



Japanese and Korean speakers' production of Japanese fricative /s/ and affricate /ts/*

Kimiko Yamakawa^{1,**} · Shigeaki Amano²

¹*Faculty of Contemporary Culture, Shokei University, Kumamoto, Japan*

²*Faculty of Human Informatics, Aichi Shukutoku University, Aichi, Japan*

Abstract

This study analyzed the pronunciations of Japanese fricative /s/ and affricate /ts/ by 24 Japanese and 40 Korean speakers using the rise and steady+decay durations of their frication part in order to clarify the characteristics of their pronunciations. Discriminant analysis revealed that Japanese speakers' /s/ and /ts/ were well classified by the acoustic boundaries defined by a discriminant function. Using this boundary, Korean speakers' production of /s/ and /ts/ was analyzed. It was found that, in Korean speakers' pronunciation, misclassification of /s/ as /ts/ was more frequent than that of /ts/ as /s/, indicating that both the /s/ and /ts/ distributions shift toward short rise and steady+decay durations. Moreover, their distributions were very similar to those of Korean fricatives and affricates. These results suggest that Korean speakers' classification error might be because of their use of Korean lax and tense fricatives to pronounce Japanese /s/, and Korean lax and tense affricates to pronounce Japanese /ts/.

Keywords: affricate, fricative, non-native speaker, production boundary

1. Introduction

The Japanese language has a voiceless alveolar fricative /s/ and voiceless alveolar affricate /ts/ (Table 1) (cf. Kubozono, 2015; Vance, 2008). These two consonants have similar spectral features: Both consist of frication noise in a frequency region higher than about 4 kHz. However, they differ in intensity envelope. /s/ tends to have a gentle onset and a long sustained duration in the intensity envelope, whereas /ts/ tends to have a steep onset and a short

sustained duration.

Unlike the Japanese language, the Korean language distinguishes a lax alveolar fricative /s/ and tense alveolar fricative /s*/ (Table 1) (Ha et al., 2009; Shin, 2015). /s/ has a longer frication duration than /s*/ at a word-initial position (Kang, 2000; Shin, 2015). /s*/ has a longer aspiration duration than /s/ at a word-initial position but it has no aspiration at a word-medial position (Shin, 2015). Additionally, the centroid of the fricative noise is lower for /s/ than for /s*/ (Cho et al., 2002). Further, in contrast to Japanese, Korean does not have an

* This study was supported by JSPS KAKENHI Grant Numbers JP21530782, JP22320081, JP26370464, and JP17K02705. We would like to thank the National Institute of Informatics, NTT Human Interface Laboratories, and Professor Hyunsoo Kim of Hongik University in Seoul for the assistance of recordings.

** jin@shokei-gakuen.ac.jp, Corresponding author

Received 24 January 2022; Revised 15 March 2022; Accepted 15 March 2022

© Copyright 2022 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Voiceless fricatives and affricates in Japanese and Korean languages related to this study

Language	Fricative		Affricate			
	Alveolar	Alveolo-palatal	Alveolar	Alveolo-palatal		
Japanese	s	ɕ	ts	tɕ		
Korean	Lax	Tense		Lax	Tense	Aspirated
	s	s*		tɕ	tɕ*	tɕ ^h

alveolar affricate /ts/. However, it does have a lax alveolo-palatal affricate /tɕ/, tense alveolo-palatal affricate /tɕ*/, and aspirated alveolo-palatal affricate /tɕ^h/ (Ha et al., 2009; Shin, 2015). At a word-initial position, /tɕ*/, /tɕ/, and /tɕ^h/ have a short, medium, and long frication duration, respectively (Shin, 2015). At a word medial position, this order of frication duration is retained. However, /tɕ/ can be pronounced as a voiced affricate between voiced sounds. In addition, /tɕ/, /tɕ^h/, and /tɕ*/ have a short, medium, and long closure duration, respectively (Shin, 2015).

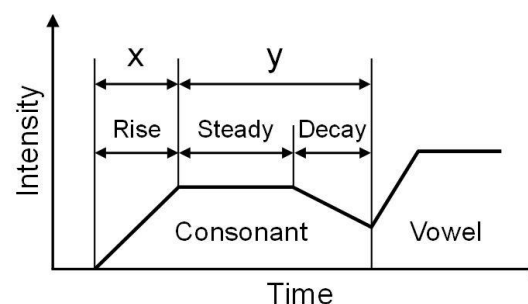
The perceptual assimilation mode for second language learners (PAM-L2) proposed by Best & Tyler (2007) predicts that speakers of native language (L1) have difficulties in discriminating phonemes in a foreign language (L2) when two phonemes in L2 are perceived to one phoneme in L1 with equal goodness or when L2 phonemes are not perceived as any L1 phoneme. According to PAM-L2 expectations, it is often observed that non-native speakers of any language have difficulty in pronouncing a foreign-language phoneme that does not exist in their first language. For example, Japanese speakers struggle to correctly pronounce English /l/ and /r/ that does not exist in Japanese (Zimmermann et al., 1984). Since the Korean language does not have an alveolar affricate /ts/ (Table 1), Korean speakers may have difficulty pronouncing Japanese /ts/. Actually, previous studies conducting a questionnaire survey for Japanese teachers on non-native Japanese learners (e.g., Matsuzaki, 1999; Sukegawa, 1993) reported that Korean speakers are not good at distinguishing Japanese /s/ and /ts/.

There are two types of pronunciation error: mispronunciation of /s/ as /ts/ (hereafter /s/→/ts/ error) and mispronunciation of /ts/ as /s/ (hereafter /ts/→/s/ error). The previous studies (e.g., Matsuzaki, 1999; Sukegawa, 1993) often reported the /ts/→/s/ error whereas rarely the /s/→/ts/ error. However, this cannot be evidence that the /s/→/ts/ error never occurs. Previous studies possibly overlooked the /s/→/ts/ error because the /ts/→/s/ error draws Japanese teachers' attention, and hence it might mask the occurrence of the /s/→/ts/ error.

Another problem of previous studies (e.g., Matsuzaki, 1999; Sukegawa, 1993) is that they mainly investigated the occurrence and characteristics of the error of /s/ and /ts/, but not the error in terms of acoustic features. The acoustic features related to the cause of Korean speakers' mispronunciation have not been clarified.

As for the acoustic features of /s/ and /ts/, Yamakawa et al. (2012) analyzed the intensity envelope of /s/ and /ts/ pronounced by native Japanese speakers and developed a method to distinguish the two consonants. They divided the intensity envelope into the rise, steady, and decay components, and then approximated these three components with lines of positive, zero, and negative slope, respectively (Figure 1). Yamakawa et al. (2012) demonstrated that /s/ and /ts/ are discriminated with a small error (1.2–6.1%) by a linear function with two variables: the rise duration and the sum of the steady and decay durations (hereafter referred to as “steady+decay”). Their results indicate that the rise and steady+decay durations are relevant

acoustic features to distinguish Japanese /s/ and /ts/. Yamakawa & Amano (2015) demonstrated that the method of Yamakawa et al. (2012) is applicable to the distinction between fricative /ɕ/ and affricate /tɕ/. Namely, they showed that these consonants are separated with low confusion errors when using the rise and steady+decay durations.

**Figure 1.** Schematic diagram of the intensity envelope and the duration of the consonant parts for analysis.

Based on these backgrounds, this study aimed to clarify the acoustic characteristics of Japanese fricative /s/ and affricate /ts/ pronounced by the introductory level of Korean learners of Japanese using the two variables (rise and steady+decay durations) proposed by Yamakawa et al. (2012). In this study, the “production boundary” is defined as an acoustic boundary obtained by discriminant analysis using the rise and steady+decay durations. The “pronunciation error” is defined as a classification error using the production boundary as a classifier. This study firstly obtained Japanese and Korean speakers' mapping of /s/ and /ts/ and their production boundaries on a coordinate plane of the rise and steady+decay durations. Then, using the Japanese speakers' production boundary, Korean speakers' pronunciation error of Japanese /s/ and /ts/ was identified to examine their characteristics. To obtain further information about the cause of pronunciation error, this study also analyzed Korean fricatives and affricates in a monosyllable pronounced by Korean speakers.

2. Speech Recording

2.1. Participants

The participants in the experiment were 24 Japanese speakers (12 males and 12 females) and 40 Korean speakers (20 males and 20 females). Their average age was 26.2 years [Minimum(Min)=21, Maximum(Max)=30, standard deviation(SD)=3.2] for the Japanese speakers, and 24.2 years (Min=19, Max=29, SD=2.6) for the Korean speakers. The Korean speakers were Japanese learners at the beginner level. They had learned the Japanese language for an average of 121 hours (Min=18, Max=300, SD=75.2) and had never lived in Japan. The participants were paid for their participation.

2.2. Word Materials

The word materials used for recording were four Japanese minimal pair words in 1–4 morae long (Table 2), having the same phoneme sequence except that their initial phoneme was /s/ or /ts/. Each item of the minimal pair words had the same accent pattern and similar auditory word familiarity (Amano & Kondo, 1999). These characteristics are desirable for a speech production experiment because the word materials are probably not affected by the difference in phoneme sequence, accent pattern, or word familiarity. The word materials served for both Japanese and Korean speakers' recordings.

Table 2. Minimal pair words with an initial phoneme /s/ or /ts/. Word meaning and auditory word familiarity are shown in parentheses and brackets, respectively

Word length (mora)	Accent pattern	Minimal pair words	
		/s/ word	/ts/ word
1	H	/suu/ (vinegar) [4.78]	/tsuu/ (harbor) [3.94]
2	LH	/suuru/ (do) [5.19]	/tsuuru/ (fish) [5.34]
3	LHL	/suneru/ (sulk) [5.66]	/tsuneru/ (pinch) [5.78]
4	LHLL	/sumagoto/ (single-string harp) [2.22]	/tsumagoto/ (multi-string harp) [2.53]

2.3. Monosyllable Materials

In addition to these word materials, Korean monosyllables with fricative or affricate consonant were used for recordings. The consonants were lax fricative /s/ (ㄱ), tense fricative /s*/ (ㄱ*), lax affricate /tɕ/ (ㄷ), and tense affricate /tɕ*/ (ㄷ*). The vowels were /u/ (ㅡ) and /ɯ/ (ㅣ). All combinations of these four consonants and two vowels yielded eight consonant-vowel type monosyllables. The monosyllable materials served only for the Korean speakers' recordings.

2.4. Procedure

The Japanese speakers participated in speech recordings in a quiet room at the National Institute of Informatics or at the NTT Human Interface Laboratories in Tokyo, Japan. The Korean speakers participated in the recordings in a quiet room at the Medialab recording studio or at Hongik University in Seoul, Korea.

For word recordings, one of the word materials was presented at the center of a computer screen in hiragana characters in each trial. Similarly, for monosyllable recordings, one of the monosyllable materials was presented in Hangul characters (i.e., one of ㅅ, ㅆ, ㅈ).

ㅅ, ㅆ, ㅈ, and ㅊ). Speakers were asked to pronounce the word or monosyllable at a normal speaking rate. Their pronunciation was digitally recorded using a microphone (ECM-999, SONY, Tokyo, Japan) and an A/D converter (UA25-EX, Roland, Hamamatsu, Japan) with 16-bit quantization and 48-kHz sampling frequency, and stored as a digital audio file on a computer. The word materials were recorded four times in a random order for each participant. Namely, there were 32 recording trials for each participant. Meanwhile, the monosyllable materials were recorded only once. Their recording order was randomized for each participant. The word materials were recorded first, and then the monosyllable materials were recorded.

3. Analysis

3.1. Japanese Speakers

The rise and steady+decay durations of /s/ and /ts/ in a word pronounced by the Japanese speakers were obtained using the estimation method proposed by Yamakawa et al. (2012). That is, the intensity envelope of frication of /s/ and /ts/ was approximated with three lines of rise, steady, and decay parts (Figure 1) by minimizing a squared error between the envelope and lines. Then, the rise and steady+decay durations were identified according to the approximated lines. Table 3 shows the mean and standard deviation of the rise and steady+decay durations of /s/ and /ts/ pronounced by Japanese speakers.

Discriminant analysis for /s/ and /ts/ was conducted using the rise and steady+decay durations as independent variables and the labels /s/ and /ts/ as the dependent variable. The discriminant function of /s/ and /ts/ for the Japanese speakers was obtained as Equation 1.

$$f = 0.038x + 0.039y - 5.414 \quad (1)$$

Where f is the predicted label, x is the rise duration (ms), and y is the steady+decay duration (ms). The discriminant error (regarded as pronunciation error) of the Japanese speakers' /s/ and /ts/ was 6.25% (Table 4). This low error ratio indicates that the discriminant analysis was successful.

By substituting zero for f in Equation 1, the production boundary of /s/ and /ts/ in the Japanese speakers was obtained as Equation 2.

$$0 = 0.038x + 0.039y - 5.414 \quad (2)$$

Figure 2 shows the Japanese speakers' /s/ and /ts/ as well as their production boundary (Equation 2). The tokens of /s/ and /ts/ are well separated by the production boundary, which corresponds to the low errors described above. The average speaking rate of the Japanese speakers was 4.52 mora/s ($SD = 0.82$ mora/s).

Table 3. M and SD of the rise and steady+decay durations (ms) of Japanese /s/ and /ts/ pronounced by Japanese and Korean speakers. The number of data is also shown

Phoneme	Japanese					Korean				
	n	Rise		Steady+decay		n	Rise		Steady+decay	
		M	SD	M	SD		M	SD	M	SD
/s/	384	76.3	33.0	101.1	30.4	640	70.8	29.7	99.9	32.3
/ts/	384	37.3	24.7	63.7	26.2	640	32.9	23.4	48.1	31.1

M, mean; SD, standard deviation.

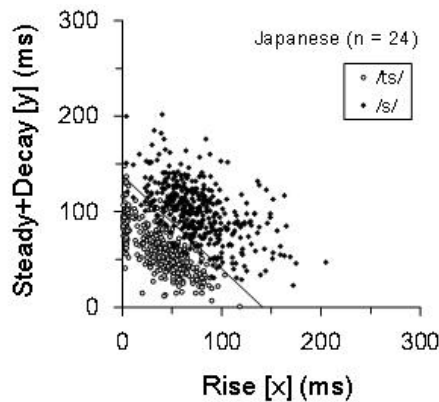


Figure 2. The Japanese speakers' /s/ and /ts/ plotted on a coordinate plane of rise and steady+decay durations. The solid line represents the production boundary of /s/ and /ts/ (Equation 2).

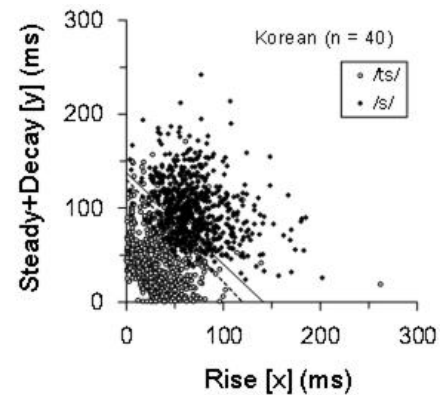


Figure 3. The Korean speakers' /s/ and /ts/ plotted on a coordinate plane of the rise and steady+decay durations. The solid and broken lines respectively represent the Japanese and Korean speakers' production boundaries (Equations 2 and 4).

3.2. Korean Speakers

As with the Japanese speakers, the rise and steady+decay durations of /s/ and /ts/ in a word pronounced by the Korean speakers were obtained using the estimation method proposed by Yamakawa et al. (2012). The mean and *SD* of the rise and steady+decay durations of /s/ and /ts/ pronounced by Korean speakers are shown in Table 3. Korean speakers' discriminant function and production boundary were obtained as Equations 3 and 4, respectively.

$$f = 0.029x + 0.027y - 3.464 \quad (3)$$

$$0 = 0.029x + 0.027y - 3.464 \quad (4)$$

where *f* is the predicted label, *x* is the rise duration (ms), and *y* is the steady+decay duration (ms). The discriminant error defined by Equation 3 was 10.47%.

Using the discriminant function of the Japanese speakers (Equation 1), the pronunciation errors of /s/ and /ts/ by the Korean speakers were identified. Namely, if *f* in Equation 1 was lower than zero for a /s/ item, it was identified as the /s/→/ts/ error, whereas if *f* was higher than zero for a /ts/ item, it was identified as a /ts/→/s/ error.

The ratios of these errors made by the Korean speakers are shown in Table 4. Their difference was tested with the z-test for two proportions. The ratio difference was significant between the Japanese and Korean speakers for the /s/→/ts/ error ($z=4.65, p<.001$) and for all errors ($z=4.63, p<.001$) but not for the /ts/→/s/ error. These results indicate that Korean speakers made more errors than Japanese speakers in the /s/→/ts/ error and all errors. The ratio difference was significant between /s/→/ts/ and /ts/→/s/ errors in the Japanese speakers ($z=2.38, p<.05$) and the Korean speakers ($z=6.89, p<.001$). These results indicate that the Japanese and Korean speakers made more errors in the /s/→/ts/ error than in the /ts/→/s/ error.

Figure 3 shows the Korean speakers' /s/ and /ts/ and the production boundary of the Japanese speakers (Equation 2) and the Korean speakers (Equation 4). Many tokens of /s/ are plotted under the Japanese production boundary, which corresponds to the high /s/→/ts/ error ratio of the Korean speakers in Table 4. The average speaking rate of the Korean speakers was 3.96 mora/s (*SD*=1.21 mora/s), which is significantly slower than the speaking rate of the Japanese speakers [$t(2,046)=10.55, p<.001$].

The rise and steady+decay durations of the Korean consonants /s/, /s*/, /tɕ/, and /tɕ*/ in a monosyllable pronounced by the Korean speakers were obtained using the estimation method proposed by Yamakawa et al. (2012). Figure 4 shows a scattergram of the consonants on a coordinate plane of the rise duration and the steady+decay duration. The production boundaries of the Japanese speakers (Equation 2) and the Korean speakers (Equation 4) were also plotted. The combined distribution of /s/ and /s*/ was very similar to that of the Japanese consonant /s/ pronounced by the Korean speakers in Figure 3. In addition, the combined distribution of /tɕ/ and /tɕ*/ had very similar distributions to that of the Japanese consonant /ts/.

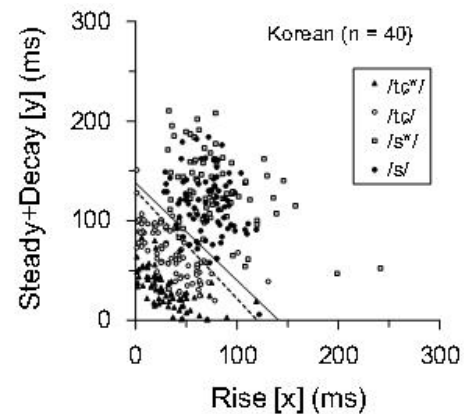


Figure 4. The Korean speakers' lax and tense fricatives (/s/ and /s*/) and lax and tense affricates (/tɕ/ and /tɕ*/) in a Korean monosyllable plotted on a coordinate plane of the rise and steady+decay durations. The solid and broken lines respectively represent the Japanese and Korean speakers' production boundaries (Equations 2 and 4) of /s/ and /ts/ at an initial position of Japanese words.

When Korean fricatives (/s/, /s*/) and affricates (/tɕ/, /tɕ*/) were categorized by the Japanese speakers' boundary of /s/ and /ts/ (Equation 2), the pronunciation error ratio was 7.5% for the fricative→affricate error, 3.8% for the affricate→fricative error, and 5.6% as a whole. Namely, as with the results in Table 4, the fricative→affricate error ratio was higher than the affricate→fricative error ratio. While, when Korean consonants /s/, /s*/, /tɕ/, and /tɕ*/ were

categorized by the Korean speakers' boundary of /s/ and /ts/ (Equation 4), the error ratio was 2.5% for the fricative→affricate error, 6.7% for the affricate→fricative error, and 4.7% as a whole. The low ratio of overall error indicated that the Korean fricatives and affricates in a monosyllable were well discriminated with the boundary derived from the Korean speakers' /s/ and /ts/ in a word.

Table 4. The Japanese and Korean speakers' pronunciation error ratio (%) of /s/ and /ts/ based on the Japanese speakers' production boundary

Speaker	Error type		
	/s/→/ts/	/ts/→/s/	All
Japanese	8.33 (384)	4.17 (384)	6.25 (768)
Korean	19.06 (640)	6.25 (640)	12.66 (1,280)

The population of the pronounced item is shown in parentheses.

4. Discussion

This study investigated Korean speakers' /s/ and /ts/ production using the rise and steady+decay durations estimated by the method proposed by Yamakawa et al. (2012), and analyzed their acoustic characteristics in terms of Japanese speakers' production boundary of /s/ and /ts/. The results indicated that Korean speakers at the beginner level make more pronunciation errors of /s/ and /ts/ than Japanese speakers, showing that Korean speakers are not good at realizing Japanese /s/ and /ts/. Their low ability in discriminating /s/ and /ts/ is reasonable because the Korean language does not have /ts/.

Furthermore, the results indicated that Korean speakers make more /s/→/ts/ errors than /ts/→/s/ errors (Table 4) which means that their pronunciation of /s/ is biased toward /ts/ that has short rise and steady+decay durations. Korean speakers' pronunciation of /ts/ is similarly biased. Their /ts/ in Figure 3 is distributed nearer to the origin of the coordinate axes than the Japanese speakers' /ts/ in Figure 2. This result indicates that Korean speakers pronounce /ts/ with shorter durations of the rise and steady+decay part than Japanese speakers do. Taken together, both distributions of /s/ and /ts/ by the Korean speakers are shifted toward the origin compared to those of the Japanese speakers. Because of this shift, the Korean speakers' production boundary (Equation 4) is located on the lower-left side (i.e., the origin side) of the Japanese speakers' production boundary (Equation 2) in Figure 3.

What causes the shift of Korean speakers' /s/ and /ts/? One possible cause for the shift of /s/ is that Korean speakers use Korean lax and tense fricatives (/s/ and /s*/) for Japanese /s/. This notion is supported by the results of /s/ and /s*/ in Korean monosyllables in Figure 4. The combined distribution of /s/ and /s*/ is similar to the distribution of Japanese /s/ in Figure 3. Furthermore, when the Japanese speakers' production boundary of /s/ and /ts/ (Equation 2) was applied for the discrimination of Korean fricatives (/s/ and /s*/) and affricates (/tɕ/ and /tɕ*/), the fricative→affricate error ratio was higher than the affricate→fricative error ratio. A similar tendency is seen in the error ratios for Korean speakers' /s/ and /ts/ in Table 4. These results suggest that Korean speakers use Korean fricatives /s/ and /s*/ to pronounce Japanese /s/. However, the /s/ and /s*/ in Figure 4 intrinsically have a different distribution from Japanese /s/ in Figure 2. Namely, they are distributed closer to /tɕ/ and /tɕ*/ than Japanese /s/ on the coordinate plane of the rise and steady+decay durations. As a result of the use of Korean fricatives having this

characteristic, Japanese /s/ pronounced by the Korean speakers may shift toward the origin.

Meanwhile, such an alternative use for Japanese /ts/ is not available because the Korean language does not have /ts/. In that case, Korean speakers probably transfer a pronunciation manner of Korean lax and tense affricates (i.e., /tɕ/ and /tɕ*/) to pronounce Japanese /ts/ because these Korean affricates have similar acoustic features to /ts/. The Korean aspirated affricate /tɕʰ/ cannot be used for the transfer because Japanese /ts/ is not aspirated by default. The idea about such transfer is supported by the result that the combined distribution of Korean /tɕ/ and /tɕ*/ in Figure 4 is similar to the distribution of Japanese /ts/ pronounced by the Korean speakers in Figure 3. Furthermore, when calculated with the Japanese speakers' production boundary, the Korean speakers showed a low affricate→fricative error rate (6.7%) similar to the /ts/→/s/ error rate (6.25%) in Table 4. These results suggest that Korean speakers transfer Korean /tɕ/ and /tɕ*/ to pronounce Japanese /ts/. Since Korean /tɕ/ and /tɕ*/ in Figure 4 intrinsically have closer distributions to the origin than Japanese /ts/ in Figure 2, the Korean speakers' transfer may result in the shift of /ts/ toward the origin.

As described above, Korean speakers pronounce Japanese /s/ and /ts/ with a shift toward the origin. As a consequence, they make more /s/→/ts/ errors and fewer /ts/→/s/ errors. Incidentally, the low /ts/→/s/ errors do not necessarily mean that the Korean speakers' /ts/ sounds natural. As seen in Figure 3, the Korean speakers' /ts/ is closer to the origin than the Japanese speakers' /ts/ in Figure 2. This means that Korean speakers pronounce /ts/ with very short rise and steady+decay durations, which correspond to the acoustic features of short duration in Korean tense affricate /tɕ*/ (Shin, 2015). The /ts/ with different acoustic features might cause an unnatural impression for Japanese speakers even though it is categorized as /ts/. This notion about naturalness should be examined in a future study.

This study is significant for instruction in Japanese pronunciation for Korean speakers because it clarifies that most Korean speakers have a common bias in pronouncing /s/ and /ts/. Korean speakers should pronounce /s/ with a longer duration of the rise and steady+decay parts. Although a /ts/ pronunciation with a short duration of the rise and steady+decay parts does not cause an error to /s/, it may degrade the naturalness and/or intelligibility of /ts/. On this point, Korean speakers also should pronounce /ts/ with a longer duration of the rise and steady+decay parts. By teaching this knowledge to Korean speakers, their pronunciation of /s/ and /ts/ will improve. They will make fewer errors between /s/ and /ts/, and will be able to pronounce /s/ and /ts/ more naturally and intelligibly.

Since the rise and steady+decay durations are time-domain variables, they probably vary with speaking rate, as would the production boundary defined by these durations. Although participants were asked to pronounce at a normal speaking rate, there might have been some speaking rate variations between participants. In particular, if the Korean speakers tended to pronounce at a faster speaking rate than the Japanese speakers, this might have caused a bias in this study because the faster speaking rate would make the duration of the rise and steady+decay shorter, causing a shift of the /s/ and /ts/ distributions toward the origin.

However, this was not the case. Korean speakers pronounced the word items at a significantly slower speaking rate than Japanese speakers as described in Section 3.2. The slower speaking rate makes the duration of the rise and steady+decay longer, and hence it cannot be the cause of the shift of /s/ and /ts/ distribution toward the

origin. If Korean speakers pronounce the word items at the same speaking rate as Japanese speakers, the /s/ and /ts/ distribution might shift closer to the origin, which would result in more /s/→/ts/ errors than the current results.

However, these notions are based on the assumption that speaking rate affects the rise and steady+decay durations. This assumption is probable, but it has not been confirmed. A future study is necessary to determine the effects of speaking rate on these durations.

This study clarified the characteristics of /s/ and /ts/ at a word-initial position, but it did not treat these consonants at a word-medial position. The /s/ and /ts/ at a word-medial position might be easily distinguished because /ts/ has a closure preceding its burst part whereas /s/ does not.

However, /ts/ at a word-medial position may have another problem for Korean speakers because Korean lax affricates at a word-medial position may appear as voiced between voiced sounds but not voiceless (Shin, 2015), although the affricates at a word-initial position always appear as voiceless. In other words, since the Korean language does not distinguish voiceless and voiced lax affricates (Ha et al., 2009), Korean speakers may mispronounce the voiceless affricate /ts/ as a voiced affricate at a word-medial position if they mimic /ts/ with a lax affricate /tɕ/. Moreover, if Korean speakers mimic the /ts/ with a tense affricate /tɕʰ/, they might mispronounce it as a geminated affricate because the tense affricate /tɕʰ/ has a long closure duration that is the main acoustic feature of the Japanese geminate affricate. For these reasons, Korean speakers might have more trouble pronouncing /ts/ at a word-medial position than at a word-initial position. Considering these points, Korean speakers' /s/ and /ts/ at a word-medial position should be examined in a future study.

One might suspect that Korean speakers' pronunciation errors are not perceived as errors by native Japanese speakers because the errors are identified with acoustical features, not perceptually in this study. Although several studies (e.g., Baese-Berk, 2019; Flege & Bohn, 2021) argued a weak or no correlation between speech production and perception, other studies (e.g., Amano & Hirata, 2010; Amano & Hirata, 2015; Denes & Pinson, 1993) claimed that speech production and perception are closely related and the production and perceptual boundaries of phonemes are expected to coincide. According to the expectation, the coincidence of the production and perceptual boundaries has been actually confirmed by experimental studies. For example, Amano & Hirata (2010; Amano & Hirata 2015) conducted an acoustic analysis and perception experiment for Japanese singleton and geminate stops at various speaking rates, and they demonstrated that production and perceptual boundaries of the stops were represented by almost the same lines on the coordinate plane of closure and subword durations. Based on these results, it is highly probable that the Japanese speakers' production and perceptual boundaries of /s/ and /ts/ are identical. If the boundaries are identical, the Korean speakers' errors identified by the production boundary should be the same as the errors identified with the perceptual boundary. Therefore, Korean speakers' pronunciation errors in this study are almost certainly perceived as errors by native Japanese speakers.

If speech production and perception have a close relationship (e.g., Amano & Hirata, 2010; Amano & Hirata, 2010; Amano & Hirata, 2015; Denes & Pinson, 1993), the perception of /s/ and /ts/ may have similar characteristics to their production observed in this

study. Namely, Korean speakers might misperceive /ts/ as /s/ more frequently than /s/ as /ts/ as a consequence of a perceptual boundary shift to the origin. This notion should be examined in a future study. However, the examination of the perceptual boundary alone is not enough to clarify the characteristics of perception of /s/ and /ts/. The sensitivity in discriminating these two consonants should also be examined because even if Korean speakers have the same perceptual boundary as Japanese speakers, they can misperceive /s/ and /ts/ because of low sensitivity in discriminating these two consonants. A future study should investigate the relationships between the perception and production of /s/ and /ts/ while paying attention to this point.

When assuming the distributions of phoneme category in production (Figures 2–4) correspond to those in perception, some implications can be provided for PAM-L2 proposed by Best & Tyler (2007). PAM-L2 distinguishes the following four cases for the perception of L1 and L2 phonological categories:

- 1) Only one L2 phonological category is perceived as equivalent (perceptually assimilated) to a given L1 phonological category,
- 2) Both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but one is perceived as being more deviant than the other,
- 3) Both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but as equally good or poor instances of that category, and
- 4) No L1-L2 phonological assimilation (Best & Tyler, 2007).

However, none of these cases fit the current relationship between Korean L1 phonemes and Japanese L2 phonemes. Korean speakers perceptually map one L2 phoneme /s/ to two L1 phonemes /s/ and /sʰ/, and they also map one L2 phoneme /ts/ to two L1 phonemes /tɕ/ and /tɕʰ/. Namely, one L2 phonological category is perceived as equivalent to two L1 phonological categories. This is a new case that PAM-L2 should consider.

Perceptual learning of Japanese L2 phonemes would occur as PAM-L2 claims, and the Korean speakers' boundary would become close to that of Japanese speakers. In that process, two Korean L1 phonemes would affect this perceptual learning. In some sense, the new case is regarded as an extension of case 1 but it is more complicated because there are two L1 phonemes and their interaction is expected. Future investigations are needed to clarify the perception and learning process in the new case, and it would improve PAM-L2.

This study showed that Korean speakers pronounce Japanese /s/ and /ts/ with a bias related to fricatives and affricates in the Korean language. Many languages such as English, Spanish, Thai, and Vietnamese do not have /ts/. Thus, /s/ and /ts/ produced by native speakers of these languages might be affected by the phonemes in their language that are similar to /s/ and /ts/, and these effects may differ according to what phonemes are similar to /s/ and /ts/. These points should be examined in a future study.

References

- Amano, S., & Hirata, Y. (2010). Perception and production boundaries between single and geminate stops in Japanese. *The Journal of the Acoustical Society of America*, 128(4), 2049-2058.
- Amano, S., & Hirata, Y. (2015). Perception and production of singleton and geminate stops in Japanese: Implications for the

theory of acoustic invariance. *Phonetica*, 72(1), 43-60.

- Amano, S., & Kondo, T. (1999). *Nihongo no goitokusei (Lexical properties of Japanese)*. Tokyo, Japan: Sanseido.
- Baese-Berk, M. M. (2019). Interactions between speech perception and production during learning of novel phonemic categories. *Attention, Perception, & Psychophysics*, 81(4), 981-1005.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Amsterdam, Netherlands: John Benjamins.
- Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193-228.
- Denes, P. B., & Pinson, E. N. (1993). *The speech chain: The physics and biology of spoken language*. New York, NY: W.H. Freeman and Company.
- Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3-83). Cambridge, UK: Cambridge University Press.
- Ha, S., Johnson, C. J., & Kuehn, D. P. (2009). Characteristics of Korean phonology: Review, tutorial, and case studies of Korean children speaking English. *Journal of Communication Disorders*, 42(3), 163-179.
- Kang, K. S. (2000). On Korean fricatives. *Speech Sciences*, 7(3), 53-68.
- Kubozono, H. (2015). Introduction to Japanese phonetics and phonology. In H. Kubozono (Ed.), *Handbook of Japanese phonetics and phonology* (pp. 1-40). Boston, MA: De Gruyter Mouton.
- Matsuzaki, H. (1999). Phonetic education of Japanese for Korean speakers. *Journal of the Phonetic Society of Japan*, 3, 26-35.
- Shin, J. (2015). Vowels and consonants. In L. Brown & J. Yeon (Eds.), *The handbook of Korean linguistics* (pp. 3-21). Malden, MA: Wiley-Blackwell.
- Sukegawa, Y. (1993). Utterance tendency of non-native Japanese speakers: Results of questionnaire survey. *Japanese speech and education, Research Report of Grant-in-Aid for Scientific Research on Priority Areas by Ministry of Education, Science and Culture*, 187-222.
- Yamakawa, K., & Amano, S. (2015). Discrimination of Japanese fricatives and affricates by production boundaries in time and spectral domains: A case study of a female native speaker. *Acoustical Science and Technology*, 36(4), 296-301.
- Yamakawa, K., Amano, S., & Itahashi, S. (2012). Variables to discriminate affricate [tʃ] and fricative [ʃ] at word initial in spoken Japanese words. *Acoustical Science and Technology*, 33(3), 154-159.
- Vance, T. J. (2008). *The sounds of Japanese*. Cambridge, UK: Cambridge University Press.
- Zimmermann, G. N., Price, P. J., & Ayusawa, T. (1984). The production of English /r/ and /l/ by two Japanese speakers differing in experience with English. *Journal of Phonetics*, 12(3), 187-193.

• **Kimiko Yamakawa**, Corresponding author
Associate Professor, Faculty of Contemporary Culture
Shokei University
2-6-78 Kuhonji, Chuo, Kumamoto, Japan

Tel: +81-96-362-2011

Email: jin@shokei-gakuen.ac.jp

Fields of interest: Speech production, L2 acquisition

• **Shigeaki Amano**

Professor, Faculty of Human Informatics

Aichi Shukutoku University

2-9 Katahira, Nagakute, Aichi, Japan

Tel: +81-561-62-4111

Email: psy@asu.aasa.ac.jp

Fields of interest: Speech perception and production