

論文

A Study of Japanese Learners' Perceptual Identification of English Fricatives

ISONO Toru

要旨

本論文の目的は、日本人学習者の英語摩擦音に対する知覚能力を実証的に検証していくことである。摩擦音に関しては、英語摩擦音の方が日本語摩擦音よりも種類が豊富なため、日本人英語学習者の特徴として、発音面においては一つの日本語音が二つの英語音に使用され（例えば、日本語の [s] が英語の [s] と [θ] の両方に用いられる等）、知覚面においてはその二つの音の判別を不得意としているという現象が広く認知されている。本研究では、日本語と英語間でそのような関係性になることが多い摩擦音において、語頭・語末そして有声音・無声音といった違いがどのように日本人学習者の英語摩擦音に対する知覚に影響を及ぼすのかを検証していく。本研究の被験者は英語を専攻している日本人大学生で、彼・彼女らにターゲット音を含む音声刺激を計200回ランダムに与え、どの英語音と認識したかを答えさせた。本論文では、その正答率を検証データとして紹介し、それをもとに日本人学習者の英語摩擦音に対する知覚能力の特徴を考察していく。

キーワード：音声知覚，摩擦音，日本人英語学習者，語頭，語末，有声音，無声音

1. Introduction

When we compare English fricatives with Japanese ones, we can firstly see that there is a difference in the total number of fricatives. The table¹ which includes both English and Japanese fricatives is shown below.

Table 1. Comparison of English and Japanese fricatives

	Labial	Coronal	Dorsal
English	[f] – [v]	[s] – [z], [ʃ] – [ʒ], [θ] – [ð]	
Japanese	[ɸ]	[s] – [z], [ʃ]	

As shown in the table, there are a larger number of English fricatives than of Japanese fricatives. This means that Japanese learners sometimes tend to substitute one Japanese consonant for more than two English consonants in their English production. It has been often discussed as the typical problems of Japanese English, since the fact that substituting one Japanese sound for more than two English sounds makes the tolerance level for inaccurate pronunciation lower when English native speakers listen to the pronunciation of Japanese learners, not the other way round. Recognising the reasons why these substitutions occur is one of the most important points of studying English production and perception by Japanese learners, so the brief description of the reasons are outlined in the rest of this section focusing on the English [f] – [v], [θ] – [ð], and [s] – [z], which are investigated in this paper.

Both the English [f] and the Japanese [ɸ] are classified as labial fricatives in the above table, and because Japanese does not have [f], the Japanese [ɸ] tends to be substituted for the English [f]. However, in a narrower analysis, they are different in that the English [f] is a labiodental fricative, which means the places of articulation are the lips and the teeth, yet the Japanese [ɸ] is a bilabial fricative, which indicates the place of articulation is only the lips, and it is pronounced by protruding the lips. The fricative noise of the English [f] is never very strong (Roach, 1991), but that of the Japanese [ɸ] is even weaker than that.

The English [v] is a voiced sound of the English [f]. However, since the corresponding voiced sound of the Japanese [ɸ] is alien to the Japanese sound systems, the Japanese [b] which is a bilabial plosive tends to be substituted for the English [v] because of the similarity in the place of articulation between the bilabial and the labiodental. Consequently, the Japanese [b] tends to be substituted for both the English [b] and [v]. In

addition to the difference in the quality of noise between the English [v] and the Japanese (the English) [b]: namely, the former has the feature of [friction] but the latter has [plosion], they are also different in the duration times. The duration of the English [v] is longer than that of the Japanese [b]. This is due to the fact that fricatives are the consonants which have the feature of [continuant], which means, to borrow Roach's (1991: 47) phrase, that "you can continue making them without interruption as long as you have enough air in your lungs", yet plosives are not continuants.

The English [θ] – [ð] are dental or inter-dental fricatives in which the tongue is placed inside the teeth or between the teeth, and since Japanese does not have any inter-dental fricatives, the Japanese [s] – [z] tend to be substituted for the English [θ] – [ð]. This would be one of the most typical substitutions characterising Japanese learners' English pronunciations, and it is often cited as the typical example supporting Lado's (1957: 27) following statement: "When a phoneme in the foreign language does not exist in the native language the student will tend to substitute the native phoneme that seems nearest within the whole structure of his native language". In addition, it has been revealed that the English [θ] – [ð] cause pronunciation problems not only for Japanese learners, but also for other language learners such as Spanish, Russian and French (Lado, 1957; Ritchie, 1968), and even for English native children (Moskowitz, 1970; Asao & Tanabe, 1993).

2. Aims of Experiment

As outlined above, the relationship between English fricatives and Japanese ones is far more complex than that between English plosives and Japanese ones, which is characterized by having the same repertory. Hence, there is fairly general agreement that English fricatives are more difficult for Japanese learners to produce than English plosives. The first aim of the experiment presented in this paper is to examine whether it is true in the case of their perception.

The second aim of the experiment is to investigate the difference in accuracy of perception between English fricatives in word-initial position and those in word-final position by Japanese learners. In the concept of markedness, features being simple and common are said to be unmarked, but those being complex and less common are said to be marked, and the latter are basically more difficult to acquire than the former in language

acquisition, because the latter often require more complex movements of the tongue muscles. In past studies, it was reported that difficulty in perception varied according to positions where target sounds appeared by Koike et al. (1977, cited in Asao and Tanabe, 1993). The subjects enrolled in that survey were around 900 Japanese learners who were studying at Japanese universities. After listening to each word including a target sound, they were instructed to listen to the other four words, and then choose one of them, which was perceived to be identical to the word previously presented. According to Asao and Tanabe (1993), despite the low percentage of the correct answer for the [b] – [v] contrast, when they were placed in word-initial position, the percentage of the correct answer was about 90%. However, conversely, in the case of the word containing [θ] in the word-final position, the accuracy rate was 95%, but it was only 51% when [θ] was placed in the word-initial position. Another study which examined the relationship between the accuracy rate in perception and positions where target sounds appear was Isono (2018). He investigated the perception of English plosives by Japanese learners, and clarified that they had greater difficulty perceiving English [p], [b] and [t] in word-final position than those in word-initial position, although such a difference was not obtained in the cases of English [d], [k], and [g]. The experiment presented in this paper sees whether a certain common feature is observed in the case of the perception of English fricatives by Japanese learners.

The concept of markedness can be applied to the contrast of voiced and voiceless sounds, and a voiced sound is generally regraded as more marked than a voiceless sound, because the former requires more complex muscle movements to vibrate the vocal folds. The third aim of the experiment is to see whether voiced English fricatives are much harder for Japanese learners to perceive than voiceless English fricatives.

In addition to the above, this paper also examines the accuracy rates on the English sounds which are often substituted by one Japanese sound in production, such as [b] & [v], [s] & [θ], and [z] & [ð], to see how much influence does the Japanese sound system have on Japanese learners' perception in English.

3. Outline of Experiment

All of the subjects taking part in Isono (2018) were also engaged in this experiment. To sum up, they were 52 university students² (16 first year students, 16 third year students,

and 20 fourth year students) majoring in English at Aichi University. The experiment was conducted using *English Accent Coach* (<https://www.englishaccentcoach.com/>), right after the experiment which was presented in Isono (2018).

This experiment gave the subjects the following two tasks. In Task 1, English words having English plosives ([p] – [b], [t] – [d], and [k] – [g]) and English fricatives ([f] – [v], [θ] – [ð], and [s] – [z]) in word-initial position were randomly given to them 100 times, and they were asked to identify the sound after they listened to each word. In Task 2, English words having the above English plosives and fricatives in word-final position were randomly given to them 100 times, and they were asked to be enrolled in the same task. In the analysis, however, it was found that the words containing [ð] in word-final position had not been given to the subjects in Task 2, so the target fricatives in the task were the above 6 plosives and the 5 fricatives ([f] – [v], [θ], and [s] – [z]) in word-final position.

In both tasks, only the above tested sounds were shown on the screen as the choices, so they had to choose their answers from the above 12 sounds.

4. Analysis

Each of the subjects gave 12 responses in Task 1, which shows the percentages of correct answers for each of the 6 fricatives and 6 plosives in word-initial position, and 11 responses in Task 2, which does the percentages of correct answers for the 5 fricatives and 6 plosives in word-final position. In total, 1196 responses (52 subjects × 23 responses) were obtained in this experiment to analyse.

4.1 Fricatives vs Plosives

The mean difference in the percentages of correct answers between fricatives and plosives is presented in the below table.

Table 2. Difference in accuracy between fricatives and plosives

	Fricatives	Plosives
Accuracy (SD)	69.25% (30.43)	95.13% (9.54)

The difference in the accuracy rates between fricatives and plosives is obvious, and the Japanese subjects had more difficulty to perceive English fricatives correctly than

English plosives, as expected. Also, as the SD values of the fricatives show, there seems to be big differences in the perceptual performance on the tested fricatives. In the next few analyses, we will see some characteristics of the perception of English fricatives by Japanese learners.

4.2 Word-initial vs Word-final

Next, we shall look closely at the difference in accuracy of perception between English fricatives in word-initial position and those in word-final position by the Japanese subjects, which is shown in the below table.

Table 3. Difference in perception of fricatives

	Word-initial position	Word-final position
Accuracy (SD)	68.51% (30.72)	70.08% (30.13)

We assumed that fricatives in word-final position would be more difficult to perceive than those in word-initial position due to the markedness features, but the accuracy of the former was a little higher than that of the latter, as the table shows. One explanation for this phenomenon is that the perception ability of English [ð] in word-final position, which might have been relatively difficult for Japanese learners, was not examined in this experiment. However, when we look at the below table which provides the breakdown of Table 3, we do not see there is much justice in that view.

Table 4. Perception of fricatives in word-initial and word-final position

	Word-initial	Word-final
[f]	88.00% (10.88)	82.09% (27.64)
[v]	49.06% (30.70)	50.00% (36.79)
[θ]	64.41% (40.11)	73.53% (28.08)
[ð]	63.86% (26.08)	——
[s]	82.23% (14.20)	80.57% (19.69)
[z]	59.28% (30.05)	63.31% (24.82)

As the figures in the above table show, significant differences in the accuracy rates between fricatives in word-initial position and those in word-final position were not observed. They were also not statistically obtained in all the cases of [f], [v], [θ], [s] and [z]: $t = 1.433$, $df = 66.44$, $p > .05$; $t = -.120$, $df=93$, $p > .05$; $t = -1.17$, $df = 58.19$, $p > .05$;

$t = .49$, $df = 92.74$, $p > .05$; and $t = -.736$, $df = 99$, $p > .05$. Therefore, the results are not consistent with the assumption that English fricatives in word-final position are more difficult to perceive than those in word-initial position for Japanese learners. In fact, it is interesting to note that the accuracy rate on the English [θ] in word-final position was much higher than that on the English [θ] in word-initial position. This is a point to which we shall return later.

4.3 Voiceless fricatives vs Voiced fricatives

In this section, we turn to the verification of the assumption that English voiceless fricatives are easier for Japanese learners to perceive than English voiced fricatives.

Table 5. Differences in accuracy between voiceless and voiced fricatives

	Voiceless fricatives	Voiced fricatives
Accuracy (SD)	76.08% (26.57)	62.18% (26.95)

As shown in the above table, the accuracy rates on English voiceless fricatives were higher than those on English voiced fricatives, and the difference was statistically significant ($T = -4.79$, $df = 343$, $p < .01$).

The below two tables show the breakdown of Table 5.

Table 6. Difference in accuracy between [θ] and [ð] in word-initial position

	[θ]	[ð]
Word-initial	64.41% (40.11)	63.86% (26.08)

Table 7. Difference in accuracy between [s] and [z] in word-initial and -final position

	[s]	[z]
Word-initial	82.23% (14.20)	59.28% (30.05)
Word-final	80.57% (19.69)	63.31% (24.82)

The difference is not statistically obtained between the accuracy rates on the English [θ] in word-initial position and those on the English [ð] in word-initial position ($t = .073$, $df = 55.24$, $p > .05$), but it is found between the accuracy rates on the English [s] and [z] in both of word-initial and -final positions ($T = 4.89$, $df = 69.23$, $p < .01$), ($T = 3.91$, $df = 101$, $p < .01$), respectively. These results suggest that the above assumption is valid in some cases, but it is not in other cases, and this is a question to be considered later.

4.4 Perception of [b] – [v], [s] – [θ], and [z] – [ð]

As noted earlier, one Japanese consonant tends to be substituted to more than two English consonants in the production of Japanese learners due to the difference in the number of consonants between Japanese and English. Concerning the sounds examined in this paper, the Japanese [b] tends to be substituted to both of the English [b] and [v], the Japanese [s] to the English [s] and [θ], and the Japanese [z] to the English [z] and [ð], and it is difficult for Japanese learners to distinguish the differences of the sounds in their English perception. In this section, we see how much the above English sounds were successfully distinguished by the Japanese subjects.

Table 8. Differences in accuracy between [b] and [v]

	[b]	[v]
Word-initial	90.09% (8.01)	49.06% (30.70)
Word-final	87.90% (20.45)	50.00% (36.79)

The above table shows that the accuracy rates on the English [b] were quite high, but those on the English [v] were about 50%, which means the probability was only one-half, in both cases of word-initial and -final positions. This result suggests that the Japanese subjects were not able to distinguish the English [v] from the English [b] successfully, and even when the English [v] was given to them, it tended to be perceived as the English [b].

The same tendency was observed in the relationship between the English [s] and the English [θ] in word-initial position as presented in the below table. The number for the English [θ] was better than a 50% chance, but a significantly big difference was obtained between the two.

Table 9. Differences in accuracy between [s] and [θ]

	[s]	[θ]
Word-initial	82.23% (14.20)	64.41% (40.11)
Word-final	80.57% (19.69)	73.53% (28.08)

However, in the case of the English [s] and [θ] in word-final position, the accuracy rates were high in both, and we see that it was relatively easy for the Japanese subjects to distinguish the English [θ] from the English [s] when they appeared in word-final position.

Finally, we shall investigate the case of the English [z] and [ð].

Table 10. Differences in accuracy between [z] and [ð]

	[z]	[ð]
Word-initial	59.28% (30.05)	63.86% (26.08)
Word-final	63.31% (24.82)	————

Unlike in the previous cases, despite the fact that Japanese has the corresponding sound [z] in its sound system, it was clarified that the Japanese subjects had difficulty in perceiving the English [z] correctly, let alone the English [ð].

In the next section, we shall examine the reasons of the phenomena which were observed in this experiment.

5. Discussion

The experiment investigated the perception of English fricatives by Japanese learners mainly focusing on the following four points: 1. The difference in the accuracy rates between English fricatives and English plosives; 2. The difference in the accuracy rates between English fricatives in word-initial position and those in word-final position; 3. The difference in the accuracy rates between voiceless English fricatives and voiced English fricatives; and 4. The effect of the Japanese sound system on Japanese learners' perception in English. In this section, we discuss the results of the last three points.

Concerning the above second point, significant differences were not observed according to the position of the English fricatives. To sum up, the accuracy rates on the English [f] and [s] were high, but those on the English [v] and [z] were relatively low, no matter whether they appeared in word-initial position or in word-final position. This result suggests that the accuracy rates on the fricatives which are easy for Japanese learners to perceive are high, but those which are relatively difficult are low, regardless of the position in which they are placed. One more point to observe is that the accuracy rates on the English [θ] in word-final position were higher than those on the English [θ] in word-initial position. In fact, the same tendency was reported by Koike et al. (1977, cited in Asao and Tanabe, 1993). The reason could be attributed to the effect of the sound following the English [θ]. Comparing the case that the English [θ] is followed by a vowel in word-initial position with the case that it is not followed by any sound in word-final position, the latter case might be easier for Japanese learners to distinguish the difference of fricative noise between the English [θ] and [s].

As for the third point, the experiment showed that the English voiced fricatives were more difficult for Japanese learners to perceive than the English voiceless fricatives. It is true that there is room for further investigation in this issue, because significant differences were not found between the English [θ] and [ð]. However, when we take the results obtained by Isono (2018) who investigated English plosives into consideration with this result, it seems reasonable to suppose that English voiced consonants are basically more difficult for Japanese learners to perceive than English voiceless consonants. A difference between the English [θ] – [ð] and the other sounds, such as English plosives and the English [s] – [z], is that the former are new sounds for Japanese learners, but the latter are not. As outlined earlier, the Japanese sound system does not have any sound corresponding to the English [θ] – [ð], but it has [p] – [b], [t] – [d], [k] – [g], and [s] – [z]. One possibility is to assume that the difference of voiceless-voiced consonants may not affect the perception of “new” L2 consonants, simply because they are difficult for learners to perceive no matter whether they are voiceless or voiced sounds.

Finally, the results of this experiments generally supported the assumption that L2 learners’ perception is conditioned by their L1 sound systems. However, we should not overlook that the accuracy rates on the English [z] were low in both of word-initial and -final positions. Although it is noted earlier that the Japanese sound system has the corresponding sound to the English [z], the conclusion that the corresponding voiced sound of the Japanese [z] is [z] is open to objection. The variation and the distribution of the Japanese [z] are quite complicated as outlined in Fujii (1999), but the characteristics summarized into the three points below would be now relatively widely accepted.

1. The Japanese [z] has two allophones [dz] and [dʒ].
2. [dz] occurs when the Japanese [z] is in word-initial position before [a], [u], [e], [o]³. On the other hand, when it is in word-medial position or is pronounced very weakly, before the vowels, it is *sometimes* realized as [z].
3. [dʒ] occurs when the Japanese [z] is in word-initial position before [i]⁴. However, when it is in word-medial position or is pronounced very weakly, before [i], it is *sometimes* realized as [ʒ].

Since [dz] and [dʒ] are perceived as much more typical Japanese sounds than [z] and [ʒ], Japanese learners of English tend to substitute the Japanese [dz] for the English [z], and the Japanese [dʒ] for the English [ʒ] (Igarashi, 1981; Imai, 1983). It seems reasonable to suppose that the above complex relationship between [z] and other sounds might have affected the result of this experiment, and it suggests that we must pay more attention to the difficulty of perceiving the English [z] by Japanese learners in educational scenes.

Notes

- 1 The framework of this table is made mainly referring to Ladefoged's (1993) viewpoint of classification.
- 2 One subject was not able to participate in the experiment presented in this paper, so the subjects were one person fewer than those who took part in Isono (2018).
- 3 Jyouo (1992) notes that the Japanese [z] is also realised as [dz] after [N].
- 4 According to Imai (1983), in the case of loanwords, when the Japanese [z] precedes [e], it is also realised as [dʒ].

References

- Asao, K. and Tanabe, Y. 1993. 「ヒアリングを構成する要素(I) 単音」. In Koike, I. (eds.). 『英語のヒアリングとその指導』. Tokyo: Taishuukan.
- Fujii, I. 1999. *Interlanguage Phonology of Japanese Speakers of English in South Wales*. Tokyo: Liber Press.
- Igarashi, S. 1981. Revised Version. *English — Its Vocal Expression*. Tokyo: Nan-un Dou.
- Imai, K. 1983. 「音声学的比較」. In Kunihiro, T. (eds.), Third Edition. 『日英語比較講座第一巻』. Tokyo: Taishuukan.
- Isono, T. 2018. Perception of English Plosives by Japanese Learners. 『言語と文化』 第40号, 1–9.
- Jyouo, H. 1992. Third Edition. 『音声学』. Tokyo: BANDAI Music Entertainment.
- Ladefoged, P. 1993. Third Edition. *A Course in Phonetics*. Fort Worth: HarcourtBrace & Company.
- Lado, R. 1957. *Linguistics across Cultures*. Ann Arbor: University of Michigan Press.
- Moskowitz, A. 1970. The Two-year-old Stage in the Acquisition of English Phonology. *Language*, 46: 426–441.
- Ritchie, W. 1968. On the Explanation of Phonic Interference. *Language Learning*, 18: 183–197.
- Roach, P. 1991. Second Edition. *English Phonetics and Phonology: a practical course*. Cambridge: Cambridge University Press.