

## 論文

# Perception of English Plosives by Japanese Learners

ISONO Toru

### 要旨

本論文では、日本人学習者の英語破裂音に対する知覚能力を実証的に検証していく。過去の研究においては、日本人が不得意とする [l] - [r] や [s] - [θ] 等の知覚能力を研究対象としたものは比較的多いが、日本語と英語間でレパートリーに差がない破裂音を取り扱ったものは少ないのが現状である。しかしながら、破裂音が語頭におかれた時には、特に無声破裂音において Voice Onset Time (VOT) や帯気音の強さにおいて日本語破裂音と英語破裂音の間には大きな違いが現れ、また語末においては、英語の有声破裂音は無声化されることが多いのに対して、日本語の場合は母音が後続するケースが多い。本研究では、そのような破裂音の日英語間の差と有標性の概念をもとに、日本人学習者が英語破裂音を知覚する時、それらが語頭にきた場合と語末にきた場合、そして無声破裂音と有声破裂音の場合では、正しく聞き取れる比率が異なってくるという仮説をたて、実験の結果を通して検証していく。被験者は英語を専攻している日本人大学生53名である。PC教室において、英語破裂音を語頭と語末に含む英単語を被験者にランダムに計200回与え、毎回どの破裂音と認識したかを答えさせた。その各英語破裂音に対する正答率が本論文における検証データとなる。

キーワード：perception (音声知覚), plosives (破裂音), Japanese learners of English (日本人英語学習者), markedness (有標性), word-initial (語頭), word-final (語末)

## 1. Introduction

The main purpose of this paper is to deal with perception of English plosives by Japanese learners who are at advanced learning stages. Although many studies have been conducted to clarify the characteristics of acquisition process of English by Japanese learners, relatively less attention has been paid to their perception ability of English sounds. In addition to the situation, many of the studies examining Japanese learners' perception focused on limited minimal pairs, such as [l] – [r], [s] – [θ], and [b] – [v] (ex: MacKain et al., 1981; Iino, 2018). This is because, as Sekiya (1994) reported, several studies (ex: Scholes, 1968; Miyawaki et al., 1975) have shown that L2 learners' perception is conditioned by their L1 sound systems, and the above English minimal pairs are typically difficult for Japanese learners since the sounds of each pair are substituted by a single Japanese sound. In this sense, it is natural that Japanese learners' perception of English plosives received scant attention, because English and Japanese plosives have the same repertory which is classified as either voiced or voiceless [p] – [b], [t] – [d], and [k] – [g]. In addition, the places of articulation are almost identical between the two languages, except for [t] – [d] (the alveolar for the English plosives, but the teeth for the Japanese ones).

## 2. Research Questions

Despite the above fact, it is also true that there are some crucial differences in the phonetic basis for the criterion between voiced and voiceless plosives in the two languages. For example, when the English voiceless plosives [p], [t], and [k] are placed in an initial position of a stressed syllable, these voiceless plosives are attended with a large amount of aspiration. This existence of a strong aspiration is one of the main ways that English native speakers distinguish the voiceless plosives [p], [t], and [k] from the corresponding voiced plosives [b], [d], and [g]. Japanese voiceless plosives are also attended with an aspiration to some degree, but, it is much weaker than that of the English voiceless plosives. This fact affects the differences in the strength of burst and VOT, and English voiceless plosives are accompanied by stronger bursts than English voiced ones and Japanese ones, and the former have much longer VOT values than the latter. In other words, Japanese voiceless plosives in word-initial position are similar to English

voiced plosives in word-initial position on these points, so we hypothesize that Japanese learners might have difficulty with English voiced plosives in their perception as a result of misperceiving English voiced plosives as voiceless ones.

Some differences are also observed when they are placed in word-final position. As is widely known, English voiced plosives in word-final position have little voicing. On the other hand, in the case of Japanese, voiced plosives in word-final position are voiced as a result of being followed by vowels. Therefore, in production, it is said that many learners have difficulty with the voice contrasts of English plosives in word-final position. This phenomenon can also be explained by the concept of markedness which is mainly approached from the perspective of language typology. Features which are simple and common in human language are said to be unmarked, but those which are complex and less common are said to be marked. In language acquisition, marked features are generally more difficult to acquire than unmarked features, mainly because marked features require more difficulty in the tongue muscles. English is a language which allows voice contrasts for all kinds of plosives [p] – [b], [t] – [d], and [k] – [g], in all word-initial, -medial, and -final positions. On the other hand, many languages including Japanese limit the occurrence of voice contrasts in word-final position (Yavas, 1994). Therefore, concerning the positions where voice contrasts occur, word-final is considered to be a more marked position than word-initial and -medial. In addition, a voiced plosive is regarded as more marked than a voiceless plosive, since the occurrence of the former in natural languages is more limited than that of the latter. For the reason given above, it can be assumed that Japanese learners especially have a difficulty with the production of English voiced plosives in word-final position, and it was clarified by Isono (2004, 2007). We will examine whether Japanese learners have difficulty with plosives in word-final position even in their perception. In past studies, Stevens et al. (1969) and Koike et al. (1978, cited in Sekiya, 1994) noted the effect of markedness on perception of foreign language sounds. In addition, the latter study reported that difficulty in perception varied according to positions where target sounds appeared.

### **3. Experiment**

#### **3.1 Subjects**

The subjects who participated in the experiment were 17 first-year students (Male =

5, Female = 12), 16 third-year students (Male = 5, Female = 11), and 20 fourth-year students (Male = 4, Female = 16). All of them were students of the department of English at Aichi University.

### 3.2 Method

The experiment was conducted in a PC room using *English Accent Coach* (<https://www.englishaccentcoach.com/>), which is a free online resource and originally designed to improve learners' ability to recognize English sounds. First, learners decide English sounds to be tested and phonetic contexts, such as the target sounds appear in word-initial or word-final position. Then, English words including the sounds in a certain condition, spoken by some different native speakers of English, are randomly given for specified number of times, and learners are asked to identify the English sounds which they heard by clicking on the phonetic symbols which are shown on the screen. Finally, scores are provided to learners and they grasp how accurately they can identify the sounds.

In this experiment, the subjects were asked to be enrolled in the following two tasks. The first task gave them English words having plosives in word-initial position randomly 100 times and asked them to identify the sounds. In the same way, English words having plosives in word-final position were randomly given to them 100 times in the second task. In both tests, only plosives were tested, so the subjects were instructed to choose their answers from the 6 choices which were [p], [b], [t], [d], [k], and [g].

## 4. Results

In this experiment, 636 responses (53 subjects × 12 responses showing the percentages of correct answers for each of the 6 plosives in word-initial and those in word-final position) were obtained.

### 4.1 Word-initial vs Word-final

The below table presents the mean difference in the percentages of correct answers between plosives in word-initial position and those in word-final position.

## Perception of English Plosives by Japanese Learners

Table 1. Accuracy in word-initial and word-final position

	Word-initial position	Word-final position
Accuracy (SD)	95.78% (7.51)	93.88% (10.74)

The accuracy rate on the plosives in word-initial position was higher than that on the plosives in word-final position, although the difference was smaller than expected.

Table 2 shows the breakdown of the above results.

Table 2. Accuracy of the plosives in word-initial and word-final position

	Word-initial	Word-final	Difference
[p]	98.52% (2.54)	93.92% (6.82)	4.60%
[b]	96.71% (3.94)	88.39% (19.7)	8.32%
[t]	98.86% (2.27)	94.49% (5.70)	4.37%
[d]	95.15% (9.63)	97.52% (2.74)	-2.37%
[k]	95.41% (5.03)	96.88% (4.72)	-1.47%
[g]	90.03% (12.1)	92.07% (12.1)	-2.04%

As shown in the table, the accuracy rates on plosives in word-initial position were significantly higher than those in word-final position in the cases of [p], [b], and [t]:  $t(52) = 4.517, p < .01$ ;  $t(52) = 2.928, p < .01$ ;  $t(52) = 5.024, p < .01$ , respectively. On the other hand, differences were not statistically obtained in the cases of [d], [k], and [g]:  $t(52) = -1.693, p > .05$ ;  $t(52) = -1.627, p > .05$ ;  $t(52) = -.909, p > .05$ , respectively. Therefore, in the cases of [p], [b], and [t], the statistical results are consistent with the hypothesis that English plosives in word-final position are more difficult to perceive than those in word-initial position for Japanese learners, but not in the cases of [d], [k], and [g].

### 4.2 Voiceless plosives vs Voiced plosives

The mean differences in the accuracy rate between voiceless plosives and voiced plosives are given in the below table.

Table 3. Differences in accuracy between voiceless and voiced plosives

	Voiceless plosives	Voiced plosives
Word-initial	97.60% (3.82)	93.96% (9.59)
Word-final	95.10% (5.92)	92.66% (13.9)

It can be seen from Table 3 that English voiced plosives were more difficult than English voiceless plosives for Japanese learners to perceive in both of word-initial and word-final positions.

The below tables provide the breakdown of Table 3.

Table 4. Difference in accuracy between [p] and [b] in word-initial position

	[p]	[b]	Difference
Word-initial	98.52% (2.54)	96.71% (3.94)	1.81%

Table 5. Difference in accuracy between [t] and [d] in word-initial position

	[t]	[d]	Difference
Word-initial	98.86% (2.27)	95.15% (9.63)	3.71%

Table 6. Difference in accuracy between [k] and [g] in word-initial position

	[k]	[g]	Difference
Word-initial	95.41% (5.03)	90.03% (12.1)	5.38%

It is clarified from the above tables that all of the English voiced plosives were more difficult than voiceless plosives for Japanese learners to perceive in word-initial position, and the values were statistically different:  $t(52) = 2.761$ ,  $p < .01$  for [p] and [b];  $t(52) = 2.718$ ,  $p < .01$  for [t] and [d]; and  $t(52) = 3.383$ ,  $p < .01$  for [k] and [g]. We see that the English voiced velar plosive [g] is especially difficult for Japanese learners.

Table 7. Difference in accuracy between [p] and [b] in word-final position

	[p]	[b]	Difference
Word-final	93.92% (6.82)	88.39% (19.7)	5.53%

Table 8. Difference in accuracy between [t] and [d] in word-final position

	[t]	[d]	Difference
Word-final	94.49% (5.70)	97.52% (2.74)	-3.03%

Table 9. Difference in accuracy between [k] and [g] in word-final position

	[k]	[g]	Difference
Word-final	96.88% (4.72)	92.07% (12.1)	4.81%

Concerning the English bilabial plosives [p] – [b] and the velar plosives [k] – [g] in word-final position, the rates on voiced plosives were lower than those on voiceless plosives, although a statistical difference was obtained only for the latter:  $t(52) = 1.894$ ,  $p > .05$  for the difference between [p] and [b];  $t(52) = 2.665$ ,  $p < .05$  for the difference between [k] and [g]. As far as the alveolar plosives in word-final position, the accuracy on the voiced plosive [d] was higher than that on the voiceless plosive [t],  $t(52) = -3.370$ ,  $p < .01$ .

Overall, the above results validate the hypothesis that English voiced plosives are more difficult than voiceless plosives to perceive for Japanese learners, except the results of the alveolar plosives [t] – [d] in word-final position.

## 5. Discussion

The first analysis verified the hypothesis that English plosives in word-final position are more difficult than those in word-initial position for Japanese learners to perceive, and clarified that the results of [p], [b], and [t] were consistent with the hypothesis, but those of [d], [k], and [g] were not. Although we cannot say for certain why differences were not observed in the cases of [d] and [k], a reason could be attributed to the low accurate rate in word-initial position as far as the case of [g] was concerned. As is generally accepted, the voiced velar plosive [g] is regarded as the most difficult sound to produce among plosives. The phenomenon that the voiced velar plosive is difficult no matter whether it is in word-initial or word-final position suggests that the above characteristic is also true in perception.

The second analysis tested the hypothesis that English voiced plosives are more difficult than voiceless plosives to perceive for Japanese learners, and it was observed in both of word-initial and -final positions, except the case of the alveolar plosives [t] – [d] in word final position. In addition to the low accuracy on the voiced velar plosive [g] which was mentioned earlier, we must pay our attention to the perceptual difficulty with the bilabial voiced plosive [b] in word-final position. As a problem concerning it, the phenomenon of mixing up it and the English voiced fricative [v] has been often reported in both of production and perception. The corresponding sound of the English [v] does not exist in the Japanese sound systems, and 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]. However, the English [v] was not included in the current experiment, so there was no chance that the low accuracy rate on the voiced English plosive [b] in word-final position was caused by the result of mixing up the two sounds. In this respect, we may say that there is another feature which makes it difficult for Japanese learners to perceive the voiced English plosive [b] in word-final position.

In this experiment, we hypothesized that English voiced plosives would be more difficult than voiceless plosives for Japanese learners to perceive, and the analysis of those in word-initial position statistically supported it. However, we should notice that the accuracy rates on the English voiced plosives [b] and [d] were sufficiently high, 96.71% for [b] and 95.15% for [d], and a problematic plosive in word-initial position was only the voiced velar plosive [g]. As far as English plosives in word-final position were concerned, this experiment suggested that they have difficulty with the voice contrast of the bilabial plosives [p] – [b] and the velar plosives [k] – [g]. In production, the difference between English voiced plosives and voiceless ones in this position is the fact that vowels preceding the former are much longer than the latter, as Roach (1991) explained. Directing learners' attention to this fact may help them to distinguish English voiced plosives from English voiceless ones in word-final position in their perception.

The results obtained in this experiment showed high accuracy rates overall, and it is natural considering the subjects' learning level and the experimental method that only plosives were given as the test items. If the subjects had been less advanced learners, the results showing more significant differences would have been obtained.

## References

- Iino, A. 2018. コンピューターを利用した高変動音素訓練 (HVPT) が日本人英語学習者が苦手とする音素の知覚に及ぼす効果, Paper presented at the 48<sup>th</sup> annual conference of CELES (the Chubu English Language Education Society).
- Isono, T. 2004. A Study of Japanese Learners' Preferred Strategies for English Plosives in Word-final Position. 『文明21』 第13号, 127–139.
- Isono, T. 2007. Interaction of the Factors Affecting the Accuracy of L2 Pronunciation—based on the acoustic analysis of English voiced plosives in word-final position. 『言語と文化』 第17号, 1–15.
- MacKain, K., Best, C., and Strange, W. 1981. Categorical perception of /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, 2, 4, 369–390.
- Miyawaki, K. et al. 1975. An Effect of Linguistic Experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics*, 18 (5): 331–340.

## Perception of English Plosives by Japanese Learners

- Roach, P. 1991. Second Edition. *English Phonetics and Phonology: a practical course*. Cambridge: Cambridge University Press.
- Scholes, R. 1968. Phonemic Interference as a Perceptual Phenomenon. *Language and Speech*, 11: 86–103.
- Sekiya, Y. 1994. 音韻の習得. In Koike, I. (eds.), 第二言語習得研究に基づく最新の英語教育. 東京 : 大修館書店.
- Stevens, K. et al. 1969. Crosslanguage study of vowel perception. *Language and Speech*, 12, 2, 1–23.
- Yavas, M. 1994. Final Stop Devoicing in Interlanguage. In Yavas, M. (eds.), 1994, *First and Second Language Phonology*. San Diego, CA: Singular.