『世界の日本語教育』12, 2002年6月

The Relationship between Production and Perception of the Stop Voicing Contrast by Korean Learners of Japanese

UEMURA Yoshiko*

Key words: **stop voicing contrast, production and perception, Korean learners, self-monitoring**

This study examined the relationship between production and perception of Japanese word-initial stops by Korean adult learners. Their perception of a native speaker's speech, as well as that of their own speech, were examined. Eighteen Korean learners produced the word-initial stop voicing contrast, and the accuracy of their production was judged by ten native speakers of Japanese. Those same learners then completed identification tests that judged their perceptual accuracy of their own pre-recorded production (internal perception), as well as their perception of a native Japanese speaker's production (external perception). The results of the comparison among their production, external perception, and internal perception showed that: 1) external perception was more accurate than internal perception; 2) perception precedes production; and 3) there was a correlation between internal perception and production. These findings indicated the importance of examining internal perception to investigate L2 learner's phonetic categorization and the relationship between perception and production. The correlation between production and internal perception that was found in this study suggests that without an accurate internal perceptual categorization to guide the sensorimotor learning of L2 sounds, the production of the L2 sounds will remain inaccurate. This study also provided pedagogical implications such as the importance of self-monitoring for phonetic acquisition.

INTRODUCTION

It is commonplace for second language (L2) learners to have readily identifiable foreign accents. Flege (1987) stated, "differences between native and non-native speakers in segmental articulation undoubtedly contribute to foreign accent" (p. 48).

* 上村佳子: Part-Time Lecturer, JET Academy and Tokyo International University, Japan Studies Program.

Korean speakers learning Japanese also have a particular foreign accent. The failure of voicing contrast in pronunciation is one of the major characteristics of Korean learners of Japanese (e.g., Bunkacho, 1971; Imada, 1989). Teachers of Japanese also know from experience that Korean learners have difficulty in acquiring the voicing contrast in Japanese. For example, Korean students tend to pronounce [kakkoo] for [gakkoo] (school) and [piiɾɯ] for [biiɾɯ] (beer). Perceptual problems caused by Korean learners in distinguishing the voicing contrast have also been pointed out by researchers (e.g., Lee, 1990) and noted anecdotally by both teachers and learners. First language (L1) negative transfer has been identified as the major reason for Korean learners' difficulty in distinguishing the voicing contrast in Japanese. This problem is caused by a cross-language difference: Korean stops have no distinctive voicing contrast, though they have three different distinctions: lax unaspirated stops, tense unaspirated stops, and aspirated stops.

Although L1 negative transfer has been well-documented, the relationship between Korean learners' pronunciation and perception of the voicing contrast in Japanese has not yet been clarified. It is not known whether learners cannot pronounce voiced stops because they cannot perceive them, or whether they cannot perceive the voicing contrast because they cannot pronounce the voiced stops. In order to investigate the relationship of their perception and production, it is necessary to determine whether their perception precedes production or vice versa, and whether there is any relationship between perception and production.

It is also meaningful to examine how Korean learners perceive their own pronunciation of the voicing contrast. Much of the research investigating the relationship between speech production and perception deals only with the perception of the speech of others (external perception), in most cases a native speaker's speech. Few studies investigate learner's self-perception of their own speech (internal perception).[1] If a "noticing of gap" plays a significant role in second language learning, as Schmidt & Frota (1986) and Schmidt (1990) argued, the self-monitoring ability of pronunciation is considered important in the development of L2 phonological acquisition as well. In this sense, it is useful to examine the learner's perception of his or her own speech, as well as the perception of the speech of others in relation to production.

Accordingly, the purpose of this study is to examine the relationship between production and perception of Japanese word-initial stops by Korean adults learning Japanese. This study will also compare Korean learners' perception of the same stops of others (external perception), with their perception of their own speech (internal perception). Following this comparison, it will examine the correlation between production and external perception and between production and internal perception. In this study both production and perception data from the same Korean subjects were collected using the same stimuli, and both aspects were also

[1] For the convenience of readers, the terminology of Sakow and McNutt (1993) will be adopted: "external perception" means one's perception of the speech of others, and "internal perception" means one's perception of one's own speech or self-perception.

judged based on the native speakers' subjective judgment.

This study was designed to answer the following questions:

1. Does mastery of the production of the stop-voicing contrast in word-initial position precede the mastery of perception, or vice versa?
2. Is external perception more accurate than internal perception in recognizing the stop-voicing contrast in word-initial position, or vice versa?
3. Is there any correlation between the abilities of production and external perception?
4. Is there any correlation between the abilities of production and internal perception?

This study will contribute to second language phonological acquisition theories in the sense that it will reveal the relationship between speech perception and production that has been an issue of debate. The examination of internal perception, which has not yet been discussed in the existing literature on Japanese stop voicing contrast, will give a new perspective on the relationship between production and perception. The study is also expected to give us pedagogical implications for teaching new L2 phones.

BACKGROUND

1 Comparison of Stops in Japanese Phonology and Korean Phonology

Japanese stops have a distinctive voicing contrast. In other words, they have voiceless stops /p, t, k/ and voiced stops /b, d, g/. Korean stops, on the other hand, have no phonemic distinction for voicing. They are all voiceless stops, but they do exhibit a three-way contrast based on aspiration and muscular tension. These contrasts are: lax unaspirated stops /p, t, k/, tense unaspirated stops /p', t', k'/ and aspirated stops /p^h, t^h, k^h/ (Shimizu, 1989).[2] Table 1 outlines the distribution of Korean stops.

Table 1 Distribution of Korean Stops

Word-Initial			Word-Medial			Word-Final
Lax	Tense	Aspirated	Lax	Tense	Aspirated	Lax
p	p'	p^h	p	p'	p^h	p
t	t'	t^h	t	t'	t^h	t
k	k'	k^h	k	k'	k^h	k

The important phonological rules relevant to this study are two intervocalic voicing rules: (a) Lax stops become voiced between voiced segments; (b) aspirated and

[2] Various kinds of terms are used to refer to Korean stops. Lax unaspirated stops and tense unaspirated stops are called, *lenis* and *fortis* (e.g., Major & Faudree, 1996) or *weak* and *strong* (e.g., Han & Weitzman, 1970). Tense unaspirated stops are also called *glottalized* stops (e.g., Lisker & Abramson, 1964). In order to avoid any confusion, the above terms, *lax unaspirated stops*, *tense unaspirated stops*, and *aspirated stops* are used in this study. Phonetic transcription of the stops also shows considerable variation.

tense stops, however, are never voiced (Martin, 1992). As a result, the predicted difficulty with stops for Korean learners in terms of negative transfer is as shown in Table 2.

Table 2 Predicted Difficulty with Stops for Korean Learners of Japanese

	Word-Initial	Word-Medial
Voiceless Stops	No difficulty	Difficulty
Voiced Stops	Difficulty	No difficulty

The predicted difficulty with stops for Korean learners of Japanese is in word-initial voiced stops and word-medial voiceless stops. Among them the word-initial voicing distinction is considered especially difficult, because there is no voiced phoneme or allophone in word-initial position in Korean. This explanation of the difficulty in terms of negative transfer agrees with the actual difficulty observed by researchers (e.g., Murasaki, 1978).

The present study does not deal with acoustic measurements to evaluate Korean subjects' production and perception abilities of stop consonants. However, it seems necessary to review the studies of stops in terms of acoustic features, especially voice onset time (VOT), which is the interval between the release of a stop closure and the onset of voicing, since it seems to be one of the most reliable acoustic cues to distinguish voiceless from voiced stops.[3] Shimizu (1989) reported the mean VOT and range of stops in Japanese and Korean. Shimizu found that there is a clear difference in VOT between the voiced stop and the voiceless stop in Japanese. According to his study, the range of Japanese /p, t, k/ was from 15 to 100 milliseconds and the range for /b, d, g/ was from −105 to −10 milliseconds. Therefore, VOT is an effective measure to distinguish the voiceless stops and the voiced stops in Japanese. Judging from his results, the VOT ranges of Japanese voiceless stops overlap with the VOT ranges of all the Korean stops. Conversely, the VOT ranges of Japanese voiced stops are completely different from the VOT ranges of any Korean stop categories. The VOT of Japanese voiced stops have negative values, though none of the Korean stops shows a negative value of VOT. Therefore, it is necessary for Korean learners to realize a new VOT category of negative values in order to acquire the Japanese voiced stop consonant.

Additionally, the phonetic distance between voiced stops in Japanese and voiceless stops in Korean may have an effect on the difficulty of Korean learners in acquiring voiced Japanese stops. Since VOT is on a continuum, the voicing contrast may fall within a "similar" sound category rather than a completely "new" category for Korean learners. Thus, it may make it more difficult for them to acquire Japanese voiced-stop sounds, according to Flege's "Equivalence Classification" hypothesis (1987, 1988) that equivalence classification prevents experienced L2 learners

[3] It is reported that there are other effective acoustic cues to distinguish the voiceless stops and the voiced stops in Japanese, such as the closure duration of stops (Sugito & Kanda, 1987).

from achieving the phonetic norm of L2 for "similar" phones but not for "new" phones.[4]

2 The Relationship Between Speech Perception and Production in L2

The interrelationship between speech perception and production has been a source of debate for decades. Generally speaking, it is widely believed that perceptual mastery precedes the acquisition of speech production based on L1 acquisition research. In L1 acquisition research, it is often said that children perceive distinctions that they cannot produce (Dresher, 1999). Furthermore, second language acquisition (SLA) researchers (Chastain, 1976; Paulston & Bruder, 1976) have also espoused the belief that perceptual mastery precedes production in their pedagogical suggestions.

This long-held belief that perception precedes production mastery was brought into question by researchers in the 1970s and 1980s. Goto (1971) and Sheldon and Strange (1982) examined the relationship between the production and perception of the English /r/ and /l/ phonemic contrast by Japanese learners, and they found that the mastery of production led over that of perception. Later, Flege has conducted various studies and written many articles regarding L2 speech production and perception (e.g., Flege, 1988, 1992, 1995, 1996; Flege & Munro, 1994). According to his Speech Learning Model (1995), "the production of an (L2) sound eventually corresponds to the properties present in its phonetic category representation" (p. 239). The model suggests that the inability to recognize or categorize perceptual distinctions limits accurate L2 production. Based on his studies so far, Flege (1996) concluded that "nonnative subjects' accuracy in producing L2 vowels and consonants (or 'sounds', for short) is related to their accuracy in perceiving L2 sounds. It appears that often, but not always, a non-native will be more accurate in perceiving than in producing L2 sounds and sentences" (p. 9).

Research into the relationship between speech production and perception has been done with other languages as well. In L2 acquisition of Spanish, Zampini (1998) examined the relationship between the production of the Spanish stop consonants /p/ and /b/ and the perception of the contrast by English-speaking learners in terms of the VOT value. However, she could not find a significant correlation between perception and production with regard to VOT.

In Japanese second language acquisition, Ogawara (1997) examined the relationship between production and perception by Korean learners of Japanese in both the segmental distinction of affricates such as /dz/ and /dʒ/ and in suprasegmental distinctions such as accent, intonation, and prominence. However, Ogawara did not find any correlation between Korean speakers' production and their perception of Japanese speakers' speech in either the segmental or suprasegmental distinctions

[4] Flege (1987) defined a "new" L2 phone as "an L2 phone which does not have a counterpart in L1, and may therefore not be judged as being the realization of an L1 category," and a "similar" L2 phone as "an L2 phone which is realized in an acoustically different manner than an easily identifiable counterpart in L1" (pp. 58–59).

above. As for the issue about which ability precedes, production or perception, he did not draw any conclusion, though in intonation and prominence the subjects seemed to perform rather well in perception, compared with the results of their production. This is probably because he used two different measures to compare production ability and perception ability. A rating was given by native Japanese speakers to measure production. On the other hand, a correctness ratio was given for performance on an identification test for perception.

Thus, there is still no clear-cut agreement about the relationship between speech production and perception. Recently, some researchers argued that the relationship should not be viewed merely as "two sides of the same coin" (Leather & James, 1996) nor as "mirror images of one another" (Major, 1998). Recently, it seems that more researchers have come to consider that no clear-cut theory exists to explain the relationship between speech production and perception for all the phonetic features.

3 Internal Perception and External Perception

Even if there is no clear correlation between production and perception of other speakers' speech, there may be some correlation between production and internal perception, because, as Borden (1980) stated, self-perception (internal perception) might appear to play an important role in forming an association between speech perception and speech production in acquiring new phonemic contrasts.

Although the number of studies that examined internal perception was small, there were some interesting findings in them. For example, the study by Sakow and McNutt (1993) was one of the studies that solely focused on the comparison of external perception and internal perception. They examined both kinds of perception of /r/ by native speakers of Japanese and Korean focusing on the judgment of the subjects' own production, and found that the subjects made more errors in internal perception in comparison with external perception. Based on these results, they suggested that learners need more training in self-monitoring.

Much earlier, in the studies of Japanese learners' perception and production of /r/ and /l/ in English by Goto (1971) and Sheldon and Strange (1982) mentioned before, the researchers examined not only learners' external perception but also their internal perception. However, their results about internal perception differed. Goto's subjects performed better in perceiving the speech of others than in perceiving their own speech, while Sheldon and Strange's subjects perceived their own speech more accurately than the speech of others.

In the area of Japanese second language acquisition, Ogawara (1997) examined the relationship between production and self-perception (internal perception) as well as the relationship between production and perception of other speakers' speech (external perception). He found that there was a significant correlation between production and internal perception, not between production and external perception, in all the distinctions he examined: affricates, accent, intonation, and prominence. He also found that learners who could use their own effective criteria in self-perception achieved higher scores on the production test. He concluded that it is not perceptual ability of other speakers' production, but monitoring ability that contrib-

utes to good pronunciation performance. He argued that it would be necessary to give learners more opportunities for self-monitoring and self-evaluation of their own pronunciation in the classroom.

Thus, the literature review showed that there is no clear explanation regarding the mechanism of speech production and perception of the stop voicing contrast for Korean learners, not to mention the relationship between production and internal perception. Therefore, it will be meaningful to examine these issues in this study.

METHOD

1 Subjects

The subjects in this study were 18 Korean students and 11 Japanese students studying at a university in the northwestern region of the United States.

The Korean subjects were enrolled in Japanese-language classes at the university. Their first language was Korean. The subjects were equally divided between males and females. They ranged in age from 19 to 32 ($M = 23$). Seven Korean students were enrolled in 1st-year Japanese at the time of the data collection, eight students in 2nd-year, one student in 3rd-year, and two students in 4th-year Japanese (see Appendix A for the profile of Korean subjects). Although their exact proficiency level was not examined, the Japanese class levels in which they were enrolled may give some indication of their level of Japanese. According to the speaking goals set in the department of the university, the majority of students who completed the 1st- and 2nd-year Japanese classes should reach the equivalent levels of Novice-High and Intermediate-Low, respectively, on the American Council on the Teaching of Foreign Languages' (ACTFL) Oral Proficiency Interview (OPI). The proficiency goals for the 3rd- and 4th-year Japanese classes should be approximately equal to Intermediate-Mid and Intermediate-High, respectively, on the OPI.[5]

All eleven Japanese subjects were native speakers of Japanese studying at the same university as above. One of them was a graduate student majoring in Japanese Literature. Her speech sample was recorded as a model native speaker's speech, and was used for the perception test. She recorded the test sentences in the Tokyo accent. The correctness of her speech in terms of voicing contrast was judged by ten other Japanese subjects. They reached 100% agreement on the distinction of the voicing contrast in her speech. The other ten Japanese subjects were undergraduate students studying at the same university: six were female and four were male. They ranged in age from 20 to 23 ($M = 22$). All of them had no Japanese teaching experience. Their length of stay in the United States was less than one year. They judged the Japanese model speaker's speech and the Korean subjects' speech.

---

[5] In this study, Korean students who differ in proficiency level — perhaps from Novice-Mid to Intermediate-High — were recruited, because it was useful to see if any pattern would emerge across all levels. However, as the number in each proficiency group could not be standardized, the effect of proficiency level on the relationship between production and perception is beyond the scope of this study.

No subjects reported any history of a speech or hearing disorder. Korean subjects were paid $5 and Japanese subjects were paid $10 for their participation.

2 Stimuli

The stimuli for both the perception and production tests consisted of 42 sentences with 32 target words and 10 filler words embedded (see Appendix B for the list of test words).[6] The sentence frame was always as follows:[7]

これは＿＿＿＿＿＿です. (This is ＿＿＿＿＿＿.)

A total of 32 test words forming 16 minimal pairs based on the word-initial stop voicing contrast were selected for test words. They included six pairs of the /p/ and /b/ contrast, four pairs of the /t/ and /d/ contrast, and six pairs of the /k/ and /g/ contrast. The following criteria were used to select the test words:

1) The stops preceding the vowels of /a/, /u/, and /e/ were chosen, because these sounds represent all the basic features for vowels: high, mid, low, front, central, and back.
2) The stop sounds /t/ and /d/ preceding /u/ change to affricates /ts/ and /dz/, respectively, in Japanese, so they were omitted.
3) The words with a /u/ sound, which potentially undergo devoicing, were also excluded.[8]
4) The number of *mora* was almost balanced among each group of the contrast.
5) The pitch accent patterns in each minimal pair were the same. The pitch accent patterns of the stimuli were consistent in each group, although all of the pair words in the /p/ and /b/ contrast have only a High/Low pitch accent pattern. This is because all words with the initial /p/ phoneme in Japanese are borrowed words or onomatopoeia, all of which contain the same High/Low accent pattern.

In addition to these 32 test words, five pairs of filler words were included, three pairs of contrasting long or short vowels and two pairs contrasting long or short

---

[6] Due to the limited availability of test words that fulfilled all the selection criteria, the test words included words that the subjects had not learned yet. However, the results of a pilot study conducted prior to the present study showed no statistically significant mean difference between the results of the test words that the subjects had learned and the test words that they had not yet learned in both production and perception tests ($p < .01$). The pilot study was conducted using almost the same method as this study; however, only six Korean subjects were used.

[7] In terms of keeping rhythm and speech rate constant, sentence stimuli is considered better than word stimuli (Ladefoged, 1993). The results of the pilot study comparing word stimuli and sentence stimuli showed no statistically significant mean difference between the results of word stimuli and sentence stimuli in all of the external perception tests, the internal perception tests, and the production tests ($p < .01$). Therefore, sentence stimuli were used in this study.

Five sentences embedded with test words may sound unnatural, because the test words in the sentence frames are verbs. This was also due to the limited availability of test words. However, in some special occasions, such as when Japanese-language teachers teach the *Kanji* for particular words, these sentences may be possible.

[8] In Japanese high vowels usually undergo devoicing when preceded and followed by voiceless obstruents.

consonants (i.e., geminate stops or single stops).

3 Procedure

3.1 Production Data Collection

Each Korean subject was given a deck of cards on which each target word was written. The cards were shuffled within the deck. The words were written in *hiragana* or *katakana*. *Kanji* letters and English translation were also written on the cards for familiarization. All the subjects had already learned and could read both *hiragana* and *katakana*. Orthographically, the voicing contrast was apparent to the subjects.

First, the subject listened to a few recorded sample sentences. Then, the subject silently read through all the target words written on the cards for familiarization. Next, in order to minimize the chance of a slip of the tongue while reading, the subject read aloud all the sentences with the target words once for practice. After that, the subject recorded the 42 target words embedded in the sentence frame with a five-second interval between them. The recording took place in a quiet recording studio with a high-quality audio tape recorder (TASCAM 202 MK II). The recording of each subject took about five minutes.

In order to evaluate the production of the Korean subjects, ten Japanese subjects listened to the edited recordings consisting of all the Korean subjects' speech together in a quiet language laboratory and were asked to judge the Korean subjects' production ability of the voicing contrast. The elicited sentences by the 18 Korean subjects were randomized using the Sound Edit Version 2 on a Macintosh computer. The Japanese subjects were asked to circle the word they heard, paying attention only to the underlined different sound in the pair. They were given three choices: two words from a minimal pair and "in-between."

The example on the identification test was as follows:

クラス　　グラス　　in-between

The subjects were asked to choose "in-between" only when they could not choose from the pair because they thought the target sound they heard was an in-between sound. This identification test took about 80 minutes in all. The subjects were given 3-minute breaks three times during the test.

3.2 Perception Data Collection

The recording made of each Korean speaker was digitally randomized so that the recording contained a mixture of the speaker's own pre-recorded sentences and those of a native speaker's model sentences. The Korean subject's speech and the native speaker's model speech were randomized every five or six sentences with the Sound Edit Version 2 on a Macintosh computer. About two weeks after the initial recording, each Korean subject was given an identification test on the edited recording. The format of the identification test was the same as the test taken by Japanese subjects. Korean subjects were told to circle the word they heard from the three choices: the two words of a minimal pair or "in-between." The subjects were asked to choose "in-between" only when they could not choose from the pair because they

thought what they heard was an in-between word in terms of the different sounds. The test of each subject took about fifteen minutes.

4 Analysis

The percentage of correctness was calculated for each subject's production performance. The average of ten Japanese judges' correctness percentage was used to indicate each Korean subject's correctness for production.[9] Since production ability should be regarded on a continuum, their "in-between" sounds were counted as 50% correct. The "in-between" sound should be regarded as an interim stage on the way to phonological mastery. The percentage of correctness of external perception was also calculated for each subject as the ratio of the number of correct answers to the number of total tokens. The correctness ratio of internal perception was calculated in the same way as external perception.

The criteria for the correct answer of each target sound in internal perception was determined by the average of the production scores marked by 10 Japanese judges. An average score 0 through 0.33 was judged as incorrect; between 0.33 and 0.66, as an "in-between" sound; and from 0.66 through 1, as correct. For example, if the average of Japanese judges' scores for a Korean subject's production of /b/ as in [*ben*] was 0.9, the subject's production of /b/ was judged as correct. Accordingly, later on if the same Korean subject perceived his or her own production of this sound as /b/ in the internal perception test, the subject's internal perception was judged as correct. However, if the subject perceived the same sound as /p/, the subject's internal perception was judged as incorrect.

RESULTS AND DISCUSSION

1 Production

The Production column in Table 3 shows 18 Korean subjects' ability to produce the voicing contrast in word-initial position indicated by the percentage of correctness. Korean subjects varied in their performance levels in their production of the voicing contrast. The range of the percentage of correctness was from 48% to 100%. The mean of their scores was 71%. Subjects 11, 16, and 17 reached 100% or more than 90% correctness in pronunciation of the word-initial voicing contrast, as judged by native speakers. On the other hand, Subjects 3, 7, and 13 made errors in almost half of the target words. Other subjects are located between them in terms of accuracy in pronunciation.

Thus, it seems that most of the Korean subjects still have a problem articulating the voicing contrast in one way or another. However, the results show that it is possible for them to acquire the contrast, since three subjects performed almost perfectly. According to Flege's Equivalence Classification Hypothesis (1987, 1988),

[9] "Correctness" in production is based on Japanese subjects' subjective judgment. In other words, it means that a Korean subject's production of a voiceless or voiced sound is located within Japanese subjects' perceptual category of each sound.

Table 3 Comparison of Production, External Perception, and Internal Perception

Subject	Production % correctness[a]	External Perception % correctness[b]	Internal Perception % correctness[c]
1	86	100	81
2	59	88	59
3	48	91	63
4	75	78	75
5	79	97	78
6	74	100	66
7	52	94	53
8	59	91	50
9	83	100	78
10	56	97	53
11	93	100	94
12	57	94	50
13	51	94	47
14	57	97	47
15	78	97	56
16	98	100	97
17	100	100	100
18	75	84	69
Mean	71	95	68
Mode	59	100	78
Median	75	97	65
Standard Deviation	16.8	6.2	17.5

[a]: Avg. No. of correct answers divided by No. of total tokens.
[b]: No. of correct answers divided by No. of tokens.
[c]: No. of correct answers divided by No. of tokens.

L2 learners hardly attain the native speakers' norm in producing L2 phones which are "similar" to their L1 phones. Since three subjects performed almost perfectly, it follows that the voicing contrast in Japanese may be interpreted as "new" to Korean subjects according to Flege's explanation. However, without acoustic measurement, it is unknown if even good speakers such as Subjects 11, 16, and 17 made the voicing contrast authentically, in terms of such acoustic features as VOT values, as Flege examined. Therefore, the result of the present study at least indicates that it is possible for Korean learners to achieve accurate pronunciation of the voicing contrast as far as the subjective judgments of native speakers are concerned.

Table 4 Error Types in Production

Subject	Number of Total Errors in Production	Number of Errors in Voiced Stops (Errors in Voiced Stops divided by Total Errors%)	Number of Errors in Voiceless Stops (Errors in Voiceless Stops divided by Total Errors%)
1	5	4 (80)	1 (20)
2	16	14 (88)	2 (12)
3	17	16 (94)	1 (6)
4	10	6 (60)	4 (40)
5	8	7 (88)	1 (12)
6	10	9 (90)	1 (10)
7	16	16 (100)	0 (0)
8	17	13 (76)	4 (24)
9	8	7 (88)	1 (12)
10	14	14 (100)	0 (0)
11	2	2 (100)	0 (0)
12	15	13 (87)	2 (13)
13	16	16 (100)	0 (0)
14	17	15 (88)	2 (12)
15	11	10 (91)	1 (9)
16	1	0 (0)	1 (100)
17	0	0 —	0 —
18	10	9 (90)	1 (10)
Total	193	171 (89)	22 (11)

Next, in order to clarify the types of errors, the ratios of errors in both voiced and voiceless stops to the total errors were calculated as shown in Table 4.

The majority of errors, or 89% of the total errors made by all the Korean speakers, was made in voiced stops, while only 11% was made in voiceless stops. "Error" here includes pronunciation that was judged as "in-between." A paired *t*-test showed a significant difference between the means of errors in voiced stops and voiceless stops ($t = 6.389$, $p < .01$). It is noteworthy that the errors made by Subjects 7, 10, 11, and 13 were all in voiced stops. This tendency supports the hypothesis based on negative transfer that Korean subjects make more errors in producing voiced stops than in producing voiceless stops in the word-initial position.

2 External Perception

The percentage of correctness of external perception was calculated for each subject as represented in the External Perception column in Table 3. Korean subjects scored

remarkably high correctness ratios on the identification test, which evaluated the ability of external perception. The mean of the ratios was high at 95% and the standard deviation was 6.2. All the subjects except Subject 4 recorded more than 80% correctness. Moreover, there were as many as 6 out of 18 subjects who recorded 100% correctness. This finding means that all the subjects except for Subject 4 had little difficulty in recognizing the voicing contrasts produced by the native Japanese speaker. One of the reasons for the high-level performances was the native speaker's very clearly articulated pronunciation. Probably the native model speaker pronounced voiced stops and voiceless stops with great distinction. Judging from the results, as far as the perception of ideal native speaker's speech production is concerned, almost all of the Korean subjects seem to have already mastered a phonetic category for the stops that are clearly articulated.

Table 5 shows the error type in external perception and demonstrates that the majority (72%) of errors in external perception appeared in voiced stops. It seems

Table 5 Error Types in External Perception

Subject	Number of Total Errors in Perception	Number of Errors in Voiced Stops (Errors in Voiced Stops divided by Total Errors%)	Number of Errors in Voiceless Stops (Errors in Voiceless Stops divided by Total Errors%)
1	0	0 —	0 —
2	4	4 (100)	0 (0)
3	3	2 (67)	1 (33)
4	7	2 (29)	5 (71)
5	1	1 (100)	0 (0)
6	0	0 —	0 —
7	2	1 (50)	1 (50)
8	3	3 (100)	0 (0)
9	0	0 —	0 —
10	1	0 (0)	1 (100)
11	0	0 —	0 —
12	2	2 (100)	0 (0)
13	2	2 (100)	0 (0)
14	1	1 (100)	0 (0)
15	1	1 (100)	0 (0)
16	0	0 —	0 —
17	0	0 —	0 —
18	5	4 (80)	1 (20)
Total	32	23 (72)	9 (28)

that there is a tendency for Korean subjects to misperceive voiced stops /b/, /d/, and /g/ produced by a native speaker as voiceless stops /p/, /t/, and /k/ more often than they misperceive voiceless stops. However, the total number of errors was so small that a significant difference was not found between the number of errors in voiced stops and voiceless stops ($t = 2.026$, $p = 0.59$).

3 Internal Perception

The Internal Perception column in Table 3 shows the Korean subjects' internal perception ability of the voicing contrast in terms of the correctness ratio. Although all of the Korean subjects performed well in perceiving the native speaker's speech, their internal perception ability varied. While Subjects 11, 16, and 17 performed almost perfectly, Subjects 7, 8, 10, 12, 13, and 14 performed rather poorly and scored around 50% correctness ratios.

The internal perception error patterns are shown in relation to production in Table 6. The patterns included how the subjects produced the target sounds as well, because it is necessary to know the subjects' intended pronunciation in order to analyze how they perceived the sounds they actually pronounced. As shown in Table 6, there are 12 error patterns in internal perception in relation to production.

The most frequent errors appeared in the pattern in which the subjects incorrectly pronounced voiced stops as voiceless stops and later perceived their own pre-

Table 6 Error Patterns in Internal Perception

Target Sound	Produced as	Perceived as	No. of Errors	(% to Total)
Voiced Stop	Voiced	Voiceless	1	(1)
		In-between	1	(1)
	In-between	Voiced	39	(21)
		Voiceless	1	(1)
	Voiceless	Voiced	95	(51)
		In-between	1	(1)
Voiceless stop	Voiceless	Voiced	29	(15)
		In-between	2	(1)
	In-between	Voiced	6	(3)
		Voiceless	3	(2)
	Voiced	Voiceless	7	(4)
		In-between	0	(0)
Total No. of Errors			186	(100)

Note: "Voiced stop" and "voiceless stop" in the first column show the target sounds that the subjects intended to pronounce. "Produced as" in the second column indicates the sounds that they actually pronounced as judged by native Japanese speakers. "Perceived as" in the third column indicates the sounds the subjects actually perceived when they listened to their own pronunciation.

recorded voiceless sounds produced by themselves as voiced sounds. This pattern of error accounted for 51% of the total errors. It is noteworthy that subjects often misperceived the sounds that were judged as voiceless stops by the native speakers almost unanimously as voiced stops. If this pattern is combined with the pattern in which subjects pronounced voiced stops inadequately as in-between and later perceived the in-between sounds as voiced sounds, this accounts for the majority (72%) of the total errors. This means that the subjects tended to mispronounce voiced stops as voiceless or in-between and later they perceived them as voiced stops. It follows that the perceptual category of voiced stops for Korean learners still partially overlaps with Japanese L1 speaker's category of voiceless stops. It seems that most of the subjects still have a different perceptual category of the voicing contrast from the native speaker's norm judging from their error patterns in internal perception.

Next, in order to see which outperformed, external perception or internal perception, the results of both are compared in Table 3. The comparison shows that all of the subjects, except Subject 17, who scored perfectly on both performances, performed better on external perception than internal perception. The mean of external perception was 95% and that of internal perception was 68%. One-way ANOVA was conducted to see if there was any significant difference between the means of external perception, internal perception, and production, and it indicated a significant difference, $F(2, 51) = 18.61$, $p < .01$. Scheffe's post-hoc test showed that there was a significant difference between the means of external perception and internal perception. This means that Korean subjects' external perception significantly outperformed their internal perception in the acquisition of the voicing contrast. This conforms with the finding of the previous studies by Goto (1971) and Sakow and McNutt (1993) that external perception precedes internal perception. It follows that Korean learners perceive the native speakers' stop consonants correctly, while they are still confused in perceiving their own stops that deviate from the native speaker's phonetic norm. It may be assumed that most of the Korean subjects could not make categorical distinctions between their own voiced and voiceless stops as native speakers of Japanese are able to do.

4 The Relationship Between Production and External Perception and Between Production and Internal Perception

In order to see which precedes — production or perception — the results of production and the results of external perception in Table 3 were compared. The one-way ANOVA and Scheffe's post-hoc comparisons showed that the mean of the Korean subjects' external perception ($M = 95$) was significantly higher than the mean of production ($M = 71$) at $p < .01$. Except Subject 17, who performed perfectly both in the production and external perception tests, all the subjects performed better in external perception than production. This finding means that almost all of the subjects could perceive the voicing contrast of the stops produced by the native speaker accurately, though many of them still had difficulty in making the stop voicing contrast. Therefore, it follows that Korean subjects' (external) perception of the word-initial voicing contrast precedes their production.

This result does not comply with the studies by Goto (1971) and Sheldon and Strange (1982), which showed that Japanese learners' production of English /r/ and /l/ preceded their perception. The finding in this study can be attributed to the fact that in the voicing contrast, the learners do not have easy access to apparent articulatory cues for the production of stops as opposed to that of English /r/ and /l/. In making the contrast of /r/ and /l/ the speaker can have tactile feedback. In other words, for example, the speaker can feel his tongue touching the alveolar ridge when he pronounces /l/. Therefore, articulatory training may be effective in improving pronunciation. This is the reason why some researchers say that the results of the study by Sheldon and Strange might be "an artifact of formal training" (Flege, 1988). However, in the distinction between voiced and voiceless stops, the speaker has no tactile feedback available. Even if the speaker tries to feel the vibration of the vocal cords by touching his or her neck when he or she produces the sound, it is impossible to realize the distinction accurately. One of the reasons for this is that the stop is not a continuous sound like vowels or approximants /r/ and /l/. Therefore, it is hard to control the motor commands during the pronunciation of it. Borden (1980) argued that "for the ballistic, fast-acting gestures of speech, such as the stop consonants, touch and audition occur too late to provide ongoing control of the motor commands" (p. 227). Another reason for difficulty in controlling the motor commands for the stop-voicing distinction may be that the distinction depends not on the articulatory control but on the timing control (the moment at which voicing starts in relation to the release of the preceding stop).

The result of the present study supports Flege's (1996) statement that often nonnatives will be more accurate in perceiving than in producing L2 sounds based on his theory that category formation exerts an influence on L2 learner's accuracy in producing L2 sounds.

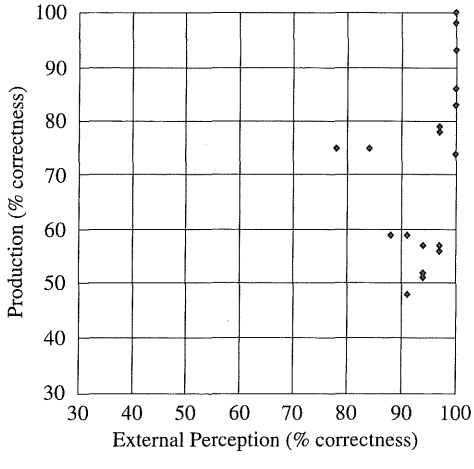

Figure 1 The relationship between external perception and production

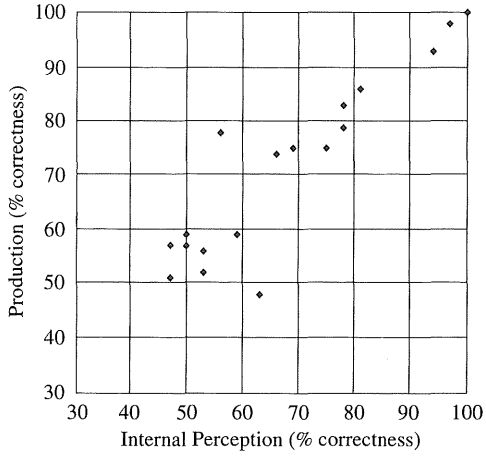

Figure 2 The relationship between internal perception and production

Next, in order to examine the correlation between production and external perception and between production and internal perception more clearly, scattered graphs were made as in Figures 1 and 2.

A Pearson correlation coefficient was obtained and indicated a significant correlation between production and internal perception ($r = .911$, $df = 16$, $p < .01$). On the other hand, there was no significant correlation between production and external perception ($r = .383$, $df = 16$, $p = .117$). Although Korean subjects performed quite well in external perception, their production ability varied. Therefore, there was no significant correlation between production and external perception. However, it was noteworthy that there was a significant correlation between production and internal perception. It follows that the better their self-perception becomes the more accurate their production will become. This result agrees with the findings that Ogawara (1997) attained in his study of Korean learners of Japanese. The findings in the current study also support Borden's (1980) claim that "self-perception [internal perception] must play an important role in forming associations between speech perception and speech production in speakers acquiring new speech patterns" (p. 238). Since production ability correlates more with internal perception than external perception, the ability of internal perception rather than that of external perception indicates the acquisition level of voicing contrast more precisely. It follows that judging only from their perceptual ability of native speaker's speech one cannot assume a L2 learner's production level.

A comparison of the results among production, external perception, and internal perception suggests insightful information concerning which contribute to L2 error production — perceptual factors or motoric (articulatory movement) factors. It is difficult to determine whether errors in L2 production can be attributed to perceptual or motoric inability. The result that external perception outperformed production of the stop voicing contrast may be misleading, because the good performance on the external perception test seems to indicate that Korean learners had almost no problem in perception. However, the correlation between internal perception and production may suggest that there was still perceptual inability that might cause errors in production, although it is unknown if motoric factors were also involved or not.

CONCLUSION

This study indicated that it is important to examine internal perception as well as external perception in order to see the relationship between speech perception and production. Korean subjects' difficulty in internal perception compared with their external perception showed that their perceptual categorization of the stop voicing contrast still deviates from the native Japanese speaker's categorization. A close look into the error patterns in internal perception revealed how Korean subjects' perceptual categories are deviated. Without examining internal perception, the incompleteness of their categorization would not have been revealed because of their satisfactory performance in external perception.

The results of the internal perception test also indicated that Korean learners' inaccurate production of the stop voicing contrast is most likely due to perceptual inability rather than some motoric inability. The finding that external perception outperforms production could not solely indicate that inaccuracy in production was attributed to perceptual inability, because almost all the Korean subjects had little difficulty in perceiving the native speaker's speech. The correlation between internal perception and production that was found in this study could imply the following causality: accurate internal perception is necessary for accurate production.

Flege claims in his Speech Learning Model that "without accurate perceptual 'targets' to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will be inaccurate" (1995, p. 238). However, Flege never distinguished between internal perception and external perception. Therefore, the results from the present study suggest that his claim requires modification to state that without an accurate internal perceptual 'target' to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will remain inaccurate.

It is noteworthy that there was not a correlation between external perception and production but between internal perception and production. This finding supports Borden's (1980) claim that internal perception plays an important role in forming associations between speech perception and speech production among speakers acquiring new phonetic contrast. Furthermore, the finding of the present study with Korean speakers suggests that internal perception rather than external perception can indicate the acquisition level of the word-initial stop voicing contrast. The correlation between production and internal perception has significant meaning in SLA phonology in the sense that it implies that internal perception reveals some aspects of L2 learners' interlanguage phonology.

The results of this study carry some pedagogical implications. The first implication is that teachers of Japanese cannot evaluate Korean learners' pronunciation ability of the voicing contrast judging only from their perceptual ability of the teacher's speech or other native speakers' speech. Even though they perform well in the identification test of the voicing contrast produced by native speakers, it does not necessarily mean that they can pronounce the contrast accurately.

Secondly, perceptual training focusing on self-monitoring might lead to accurate production. The result of this study suggests an importance of self-monitoring. As internal perception correlates with production, there is a great possibility that an improvement in internal perception leads to an improvement in production. L2 learners need more opportunities for self-evaluation of their own speech and more training in self-monitoring of their own production. The self-monitoring also gives L2 learners a chance of consciousness raising for learning pronunciation that is advocated by Celce-Murcia, Brinton, and Goodwin (1996). Without recognizing their own production and the errors they make, L2 learners cannot correct them.

There are several suggestions for further study. First, one of the future directions of this study will be one that encompasses an acoustic approach. It is assumed that acoustic measurements such as VOT values and the onset of sounds, and the use of synthesized speech will reveal Korean learners' acoustic features of production and

the perceptual boundary more closely. Furthermore, to gauge the phonetic distance between Japanese voiced stops and each category of Korean voiceless stops perceived by Korean speakers is regarded as necessary to predict the difficulty of phonetic categorization, according to Flege's Speech Learning Model. It is meaningful to examine how Korean learners judge the similarity of each Japanese stop in terms of their L1 stops. Another direction for future study will naturally be one that examines how the training of self-monitoring affects speech production to prove the significance of self-monitoring for speech learning. If the efficacy of training students to self-monitor is proved, it will indicate a causal relationship between internal perception and production.

BIBLIOGRAPHY

Borden, G.J. 1980. Use of feedback in established and developing speech. In N. Lass (Ed.), *Speech and language* (Vol. 3, pp. 168–93). New York: Academic Press.

Celce-Murcia, M., Brinton, D.M., & Goodwin, J.M. 1996. *Teaching pronunciation: A reference for teachers of English to speakers of other languages.* New York: Cambridge University Press.

Chastain, K. 1976. *Developing second language skills: Theory to practice.* Chicago: Rand McNally.

Dresher, B.E. 1999. Child phonology, learnability, and phonological theory. In W.C. Ritchie, & T. Bhatia (eds.), *Handbook of child language acquisition* (pp. 317–46). California: Academic Press.

Flege, J.E. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15,* 47–65.

———. 1988. The production and perception of foreign language speech sounds. In H. Winitz (ed.), *Human communication and its disorders: A review 1988, 2.* Norwood, NJ: Ablex Publishing Corp.

———. 1992. Speech learning in a second language. In C.A. Ferguson, L. Menn, & C. Soel-Gammon (eds.), *Phonological development: Models, research, implications.* Timonium, Maryland: York Press.

———. 1995. Second-language speech learning: Theory, findings, and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, Maryland: York Press.

———. 1996. *Comparing L2 production and perception: What level of analysis?* Paper presented at the third joint meeting of the Acoustic Society of America and the Acoustic Society of Japan, Honolulu, HI.

Flege, J.E. & Munro, M.J. 1994. The word unit in second language speech production and perception. *Studies in Second Language Acquisition, 16,* 381–411.

Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia, 9,* 317–23.

Han, M. & Weitzman, R. 1970. Acoustic Features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica, 22,* 112–28.

Ladefoged, P. 1993. *A course in phonetics* (3rd ed.). Fortworth, TX: Horcourt Brace Javanovich.

Leather, J. & James, A. 1996. Second Language Speech. In W. Ritchie & T. Bhatia (eds.), *Handbook of second language acquisition* (pp. 269–316). San Diego, CA: Academic Press.

Lisker, L. & Abramson, A. 1964. A cross-language study of voicing in initial stops: Acoustical

measurements. *Word, 20*, 384–422.
———. 1972. Voicing timing in Korean stops. *Proceedings of the Seventh International Congress of Phonetic Sciences* (pp. 439–46). The Hague: Monton.
Major, R. 1998. Interlanguage phonetics and phonology: An introduction. *Studies in Second Language Acquisition, 20*, 131–37.
Major, R. & Faudree, M. 1996. Markedness universal and the acquisition of voicing contrast by Korean speakers of English. *Studies in Second Language Acquisition, 18*, 69–90.
Martin, S.E. 1992. *A reference grammar of Korean: A complete guide to the grammar and history of the Korean language*. VT: Charles E. Tuttle.
Paulston, C.B. & Bruder, M.M. 1976. *Teaching English as a second language: Techniques and procedures*. Cambridge: Winthrop Publishers, Inc.
Sakow, M. & McNutt, J.C. 1993. Perception of /R/ by native speakers of Japanese and Korean: Internal and external perception. *International Review of Applied Linguistics 31*, 46–53.
Schmidt, R. 1990. The role of consciousness in second language learning. *Applied Linguistics, 11*. 129–58.
Schmidt, R. & Frota, S. 1986. Developing basic conversational ability in a second language: A case-study of an adult learner of Portuguese. In R. Day (ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Newbury House.
Sheldon, A. & Strange, W. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics 3*, 243–61.
Shimizu, K. 1989. A cross-language study of voicing contrast of stops. *Onsei kagaku kenkyu [Studia Phonologica], 23*, 1–12.
Zampini, M.L. 1998. The relationship between the production and perception of L2 Spanish Stop. *Texas Papers in Foreign Language Education, 3* (3), 85–100.

今田滋子（1989）『発音』国際交流基金　教師用日本語教育ハンドブック第6巻，凡人社.
小河原義朗（1997）「発音矯正場面における学習者の発音と聞きとりの関係について」，『日本語教育』第92号，日本語教育学会
杉藤美代子・神田靖子（1987）「日本語と中国語話者の発話による日本語の無声および有声破裂子音の音響的特徴」，『大阪樟蔭女子大学論集』24号.
福岡昌子（1995）「北京語・上海語を母語とする日本語学習者の有声・無声破裂音の横断的および縦断的習得研究」，『日本語教育』第87号，日本語教育学会.
文 化 庁（1971）『音声と音声教育』，大蔵省印刷局.
村崎恭子（1978）「アジア留学生の日本語のクセ」，『言語生活』第322号.
李 炯 宰（1990）「韓国人の日本語学習者の音声教育に関する研究——発音および聞き取り上の問題点を中心に」，『日本語と日本文学』第12号，筑波大学国語国文学会.

Appendix A

Profile of Korean Subjects

Subject	Male/ Female	Age	Enrollment in Japanese Class	Other experiences of learning Japanese	Length of stay in the United States
1	F	22	1st-year class		1 year
2	F	20	1st-year class		5 years
3	F	20	1st-year class		6 years
4	M	27	1st-year class		3 years
5	F	19	1st-year class		5 years
6	M	32	1st-year class	1 month self-study	7 years
7	M	20	1st-year class		3 years
8	M	26	2nd-year class		2 years
9	F	22	2nd-year class		2 years
10	F	20	2nd-year class	2 years in high school	4 years
11	M	21	2nd-year class		7 years
12	F	21	2nd-year class		3 years
13	F	20	2nd-year class	6 months in high school	5 years
14	M	28	2nd-year class	1 year in high school	2 years
15	M	20	2nd-year class	1 year in high school	7 years
16	M	27	3rd-year class	3 years in Japan	7 months
17	F	26	4th-year class		4 years
18	M	27	4th-year class		5 years

Appendix B

Minimal Pairs Used as Stimulus Materials

32 Target Words:

(translation)

Group1: /p/ & /b/

パン [pan] (bread)	ばん [ban] (turn)
パット [patto] (putt)	バット [batto] (bat)
プレーン [pɯreen] (plain)	ブレーン [bɯreen] (advising intellectuals)
プースケ [pɯɯsɯke] (imaginary name)	ブースケ [bɯɯsɯke] (imaginary name)
ペン [pen] (pen)	べん [ben] (valve)
ペンチ [pentʃi] (plier)	ベンチ [bentʃi] (bench)

Group2: /t/ & /d/

たいがく [taigakɯ] (withdrawal from school)	だいがく [daigakɯ] (university)
たく [takɯ] (boil)	だく [dakɯ] (embrace)
てんき [tenki] (weather)	でんき [denki] (electricity)
てる [teɾɯ] (shine)	でる [deɾɯ] (go out)

Group 3: /k/ & /g/

かくせい [kakɯsee] (different age)	がくせい [gakɯsee] (student)
かむ [kamɯ] (chew)	ガム [gamɯ] (chewing gum)
クラス [kɯɾasɯ] (class)	グラス [gɯɾasɯ] (glass)
くうぜん [kɯɯzen] (unprecedented)	ぐうぜん [gɯɯzen] (accident)
けた [keta] (figure)	げた [geta] (wooden clogs)
けんり [kenɾi] (right)	げんり [genɾi] (principle)

10 Filler words:

おじさん [odʒisan] (uncle)	おじいさん [odʒiisan] (grandfather)
ちず [tʃidzɯ] (map)	チーズ [tʃiidzɯ] (cheese)
かど [kado] (corner)	カード [kaado] (card)
じかん [dʒikan] (time)	じっかん [dʒikkan] (realization)
にし [niʃi] (west)	にっし [niʃʃi] (diary)