名古屋学院大学論集　言語・文化篇　第 22 巻　第 1 号（2010 年 10 月）

Acoustic Analysis of English and Japanese Stop Voicing Contrasts Produced by Korean L2 Learners

Katsumasa Shimizu

Abstract

The present study examines phonetic characteristics of Korean L2 learners, when they study English and Japanese stop voicing contrasts. VOT was measured for their Korean, English and Japanese utterances, and it was found that VOT can distinguish a three-way contrast of Korean stops, and that Korean L2 learners use their tense, unaspirated stops for voiced stops in English and Japanese, but it was not clearly shown how they use other types of categories for voiceless stops in English and Japanese, though they seem to use a strongly aspirated stop for English voiceless stops. Phonological features /VOICE, ASPIRATED, TENSE/ were proposed to specify the laryngeal characteristics of Korean, English and Japanese.

Index terms: Korean L2 learners, VOT, stop voicing, English, Japanese

1. Introduction

This paper examines phonetic characteristics which Korean L2 learners exhibit in the process of acquiring stop voicing contrasts in English and Japanese. Korean is known to have a three-way contrast in its stop consonants, and they differ in manner and point of articulation. They are usually classified as voiceless, tense, unaspirated stops (Type 1 stops), voiceless lax, weakly aspirated stops (Type 2 stops), and voiceless, strongly aspirated stops (Type 3 stops).[1] Some phoneticians simply call them as fortis, lax, and aspirated stops, respectively, while others call them strong, weak and aspirated. The disagreements on the terms to apply to the stops reflect some difficulty in characterizing their phonetic features on purely phonetic grounds. There are no voiced stops in their phonemic inventory, except in a few cases.[2]

The contrast in Korean stops has been of interest among phoneticians, since it involves several modes of vocal fold adjustments, and has been studied from physiological, acoustic and phonological points of view. First, physiologically, Kim (1965) carried out an electromyographic test of bilabial stops and reported greater muscle activity of the lips for Type 1 stops. Hirose et al. (1974) examined the activities of the intrinsic laryngeal muscles and found that these muscles serve to differentiate the three types of stops, and that Type 1 stops are specifically characterized by the activity of the

thyroarytenoid muscle. Further, Kagaya (1974) examined the glottal width and the timing of the closure by using a fiberscope, and reported that there is a considerable difference among the three types of stops in glottal width; the smallest glottal width for Type 1 stops, the intermediate for Type 2 stops, the largest for Type 3 stops. Next, acoustically, there have been several studies, and Han and Weitzman (1970) and Shimizu (1996) examined such features as voice onset time, Fo curve and other acoustic characteristics for distinguishing the three types of stops, and reported that VOT and Fo curve are both functional for characterizing the three-way contrasts. Furthermore, aerodynamically, Dart (1987) measured the air pressure and oral flow of Type 1 and Type 2 stops, and found that Type 1 stops are produced with higher intra-oral pressure before the release, but a lower flow after the release. Through these studies, it can be stated that the three stops in Korean differ in laryngeal muscle activity, glottal width, and timing of glottal closing relative to articulatory release and these differences in articulation are reflected in such features as VOT, Fo, F1 and intensity (Shimizu, 1996). The problem is the lack of examination on how these features are coordinated with each other, and is how these features are reflected when Korean L2 learners acquire the two-way contrast of stop consonants in English and Japanese as a second language.

The present study was undertaken as part of a research project to examine voiced–voicing contrasts of stop consonants, and was aimed at examining how Korean L2 learners exhibit the phonetic characteristics in learning English and Japanese stop voicing. Japanese and English have two types of voicing contrast for stops: voiced /b, d, g/ and voiceless /p, t, k/, and are two-category languages in the classification of voicing contrast. The difference between the two categories in Japanese is truly one of voicing, but the situation in English is a bit different from those in Japanese. Voiced stops in English are fully voiced in the middle position, while they tend to be voiceless unaspirated in the initial one. Voiceless stops in English are aspirated in the initial position but are unaspirated in other positions. As is generally known, the stops in English occur in the initial, medial and final positions, while those in Japanese do in the initial and medial positions, and they are contextually influenced by phonetic environments. It will be significant to examine how a three-way contrast of stops in Korean is correlated to a two-way contrast in English and Japanese, since it will clarify the laryngeal relationship in languages, and the study will reveal how the three-way contrast of Korean stops is manifested in a two-way contrast in the acquisition of English and Japanese as a second language.

2. Experimental Procedure

2.1 Subjects

The subjects in the present study are four native speakers of Korean, three males and one female, and are speakers of the Seoul dialect. They are undergraduate students of Osaka University, and are studying Japanese as a foreign language. They have studied English for eight years in Korea. Their

ages are 22–23 years.[3] They have resided in Japan for a period of 1 to 2 years, and their Japanese proficiency is in advanced level, while their English proficiency is in intermediate level.

2.2 Linguistic Materials

Linguistic materials for recording in the present study are shown below. They include six near-minimal triplets, differing in the initial stops, nine minimal pairs in English and six minimal pairs in Japanese. They were written in Korean characters and were presented to each subject. These linguistic materials were read twice by each subject. The recording was made in a sound-proof recording room at Osaka University.

Korean Linguistic Data

Type 1

[p*ul]	horn	[pul]	fire	[p^{h}ul]	grass
[t*am]	sweat	[tam]	wall	[t^{h}am]	envy
		[koŋ]	zero	[k^{h}oŋ]	bean
[k*ul]	honey	[kul]	oyster		

[p*aruda]	be quick				
[t*al]	daughter	[t*am]	sweat	[t*uda]	float
[k*ul]	honey	[k*um]	dream	[k*i:da]	place
[pada]	sea				
[tal]	moon	[to:l]	stone		
[kat]	hat	[kot]	place	[kilda]	be long
[p^{h}al]	arm	[p^{h}i]	blood		
[t^{h}a:l]	a mask	[t^{h}ək]	chin	[tot^{h}ori]	acorn

English Linguistic Data

1. It's cold.	It's gold.
2. Did she have the time?	Did you have the dime?
3. Her curls are good.	Her girls are thirteen years old.
4. We say pig.	We say big.
5. We say tick.	We say Dick.
6. We say kill.	We say gill.
7. We say (back pack) again.	
8. We say (tack dark) again	
9. We say (cat gat) again.	

Japanese Linguistic Data

これは　（　　ピン　　）です。(This is a pin.)

これは（　瓶　ビン　）です。(This is a bin.)
これは（　点　　）です。(This is a point.)
これは（　田　デン　）です。(This is a rice field.)
これは（　犬　　）です。(This is a dog.)
これは（　原　ゲン　）です。(This is a field.)
これは（　パン　　）です。(This is bread.)
これは（　板　バン　）です。(This is a board.)
これは（　単　タン　）です。(This is a unit.)
これは（　団　ダン　）です。(This is a group.)
これは（　缶　　）です。(This is a can.)
これは（　雁（ガン）　）です。(This is a wild goose.)

2.3 Acoustic Analysis

Acoustic analysis of the recorded materials was made through Acoustic Core 8 of Arcadia Co. The linguistic materials were digitized and analyzed at a sampling rate of 44.1 kHz, and the digitized materials were stored for reviewing and listening. The measurement of durations such as voice onset time was made by manually positioning two cursors in the display of the waveform and wide-band spectrograms. Figure 1 shows the three-category of Korean stops [p*ul] – [pul] – [p^hul].

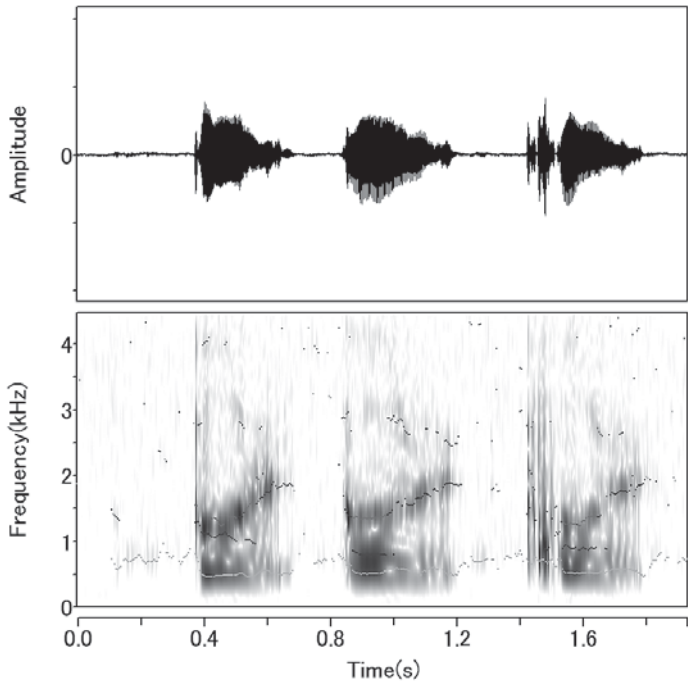

Figure 1　Waveforms and Spectrograms of [p*ul](left)–[pul] (middle)–[p^hul] (right)

3. Results

The measurement of voice onset time (VOT) was made for the interval between the consonant release and the onset of voicing for three types of Korean stops, and for two types of English and Japanese stops. Table 1 indicates the mean VOT value (ms) for the three types of Korean stops.

Table 1 *Mean VOT values of Korean Stops (ms) (N = 24 for Type 1 and 2 stops, N = 12 for Type 3 stops) (s.d. in parentheses)*

Type 1 stops		Type 2 stops		Type 3 stops	
p*	19.0(5.7)	p	60.7(21.8)	p^h	95.4(20.2)
t*	17.8(9.6)	t	61.3(14.9)	t^h	90.0(7.3)
k*	33.6(12.1)	k	66.9(23.2)	k^h	90.3(7.6)

It can be seen from Table 1 that VOT values increase in the order from Type 1 stops to Type 3 stops, and there is a difference among three categories of Korean stops. It is apparent that Type 3 stops are produced with considerable delay of voicing and are strongly aspirated. The result and tendencies are in general agreement with previous studies (Hardcastle, 1973, Shimizu, 1996). In Shimizu (1996), Type 1 stops showed the shortest value of VOT, Type 2 stops had an intermediate value between Type 1 and Type 3 stops, and Type 3 stops showed a characteristically longer value of VOT than those of other two types. Although there is a tendency that velar stops show longer delay of voicing among the three places of articulation, this was not clearly observed in the three types of stops, though Type 2 stops seemed to show this trend slightly.

Table 2 shows the mean VOT values of English stops produced by Korean speakers. From Table 2, it can be said that voiced stops in English are produced with some voicing delay, and that voiceless stops in English are produced with a long delay of voicing, and the VOT values are in the range of Type 3 stops in Korean. It can also be found that voiced stops are articulated in the range of+VOT area, which can be considered rather unusual.

Table 2 *Mean VOT values of English Stops by Korean speakers (ms) (N = 16) (s.d. in parentheses)*

Voiced		Voiceless	
b	24.3(14.3)	p	79.1(30.8)
d	13.8(28.7)	t	70.5(20.9)
g	39.3(16.3)	k	101.0(27.5)

Next, Table 3 shows the mean VOT values of Japanese stops uttered by Korean L2 learners, and it can be said that the two categories are clearly distinguished in the VOT dimension, and voiced stops in Japanese are produced with a voicing delay, as found in the case of English voiced stops.

Table 3 *Mean VOT values of Japanese Stops by Korean speakers (ms) (N = 12) (s.d. in parentheses)*

Voiced		Voiceless	
b	19.5(9.0)	p	71.8(22.7)
d	23.5(12.0)	t	63.0(18.1)
g	34.6(8.7)	k	72.5(26.7)

It can be said that Korean L2 learners show a voicing delay in the production of English and Japanese voiced stops, though they were supposed to show voicing lead which is generally found in the production of voiced stops in the two languages.

4. Discussion

The examination of voice onset time (VOT) reveals that a three-way stop in Korean can be distinguished by the VOT dimension, i.e., three domains such as short range, medium range and long range on the scale of VOT values. As is generally known, Type 1 stops are called voiceless tense unaspirated stops, Type 2 stops are called voiceless lax unaspirated stops, and Type 3 stops are called voiceless (tense) aspirated stops. From Table 1, it can be seen that the lax unaspirated stops are weakly aspirated in word-initial position. Besides VOT values, it is known that other features such as Fo and Fo contour immediately after the release of closure, and F1 values may be relevant to the distinction, and it can be said that such a distinction involving tensing of laryngeal muscles can be manifested in several acoustic features.[4]

In comparing the VOT values of Korean stops with those of English and Japanese stops in Tables 2 and 3, it can be said that Korean L2 learners closely align their Type 1 stops with English and Japanese voiced stops, but they do not show consistent alignment of their voiced stops with English and Japanese voiceless ones, as shown in Figure 2.

From Figure 2, it can be found that Korean L2 learners use Type 2 stops for voiceless stops in Japanese, since the VOT values of Type 2 stops are close to those for Japanese voiceless stops. Since Korean lacks a voiced/voiceless contrast, their L2 learners tend to use the categories which are close to their own VOT values. In examining the results in Figure 2, it can be said that strongly aspirated stops in Korean roughly correspond to the ones of voiceless aspirated stops in English, and lightly aspirated and lax stops do voiceless stops in Japanese.

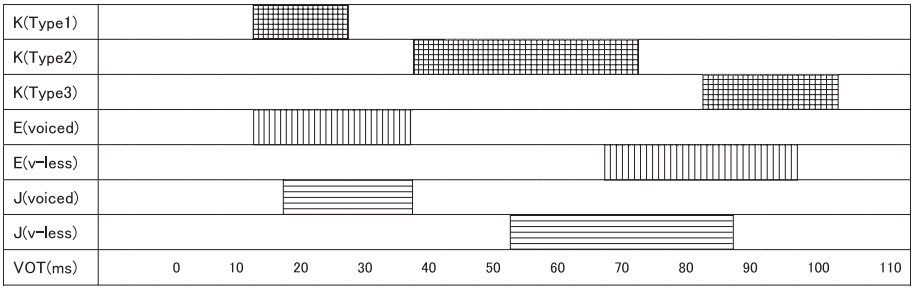

Figure 2 VOT Ranges of Korean, English, and Japanese Stops Produced by Korean L2 Learners
K(Type1) = Korean tense unaspirated stop, K(Type2) = Korean lax unaspirated stop
K(Type3) = Korean strongly aspirated stop
E(voiced) = English voiced stops, E(v-less) = English voiceless stops
J(voiced) = Japanese voiced stops, J(v-less) = Japanese voiceless stops

Several studies have been done on VOT values in English and Japanese stops produced by the natives speakers of both languages, and according to Shimizu (1999), the values can be shown as follows:[5]

Table 4 *Mean VOT values of Japanese and English Stops (ms) (s.d. in parentheses)*

	Japanese	English
/p/	41(17.1)	68(15.3)
/t/	30(12.7)	82(18.6)
/k/	66(12.1)	85(20.1)
/b/	−89(28.5)	−88(18.1)
/d/	−75(32.7)	−74(28.0)
/g/	−75(27.0)	−85(14.4)

(From Shimizu (1999))

As shown in Table 4, there is a clear-cut distinction in VOT values for voiced and voiceless stops in both Japanese and English, and it can be noted that voiceless stops in Japanese show a medium length of voicing lag, while those in English show a medium to long voicing lag. Since voicing lag is related to aspiration, it can be said that voiceless Japanese stops are moderately aspirated, while those in English are strongly aspirated. In connection with this, it can be noted that the VOT values for the voiceless velar stops in both languages have a greater value than those of bilabial and alveolar stops, and this is in accordance with previous studies.[6] It is known that the VOT value increases as the point of articulation goes backward in the oral cavity, but this is not observed between bilabial stops and alveolar stops in Japanese. For the voiced stops there is a considerable voicing lead before the release of consonant closure, and there doesn't seem to be any correlation between its value and the point of

articulation.

5. Phonological Features for Stop Consonants in Three Languages

There has been a great deal of discussion on laryngeal features in phonology; what are the features for, what are the physical bases for the features, and what are the conventions in phonological descriptions, and so on. The feature is a basic unit in phonology and is considered to show the relation among the segments. In order to deal with cross-language voicing contrasts, it is necessary to have three phonological features, such as /VOICE/, /ASPIRATED/ and /TENSE/.

The features are considered to be abstract in nature and are to have binary values of plus and minus. It is assumed that other specific features can be added to the list of the features to make a more accurate description, if necessary. For the present study, three features are sufficient. The stops /p, t, k/ in Japanese are specified as [-voice], and the same stops in Korean are specified as [-voice, -aspirated, -tense]. This means that voicing contrasts in Japanese involve one feature, whereas those in Korean are different from Japanese because they require more feature specifications.

The phonemic inventory of stops in the three languages can be shown as follow:

Japanese	/b/ - /p/	/d/ - /t/	/g/ - /k/
English	/b/ - /p/	/d/ - /t/	/g/ - /k/
Korean	/p*/ - /p/	/t*/ - /t/	/k*/ - /k/
	$/p^h/$	$/t^h/$	$/k^h/$

Although Japanese and English use two series of stops, they differ in physical features, and as shown in Table 4, English /p, t, k/ show a long voicing delay, which means a strong aspiration, but Japanese /p, t, k/ show a medium voicing delay. As shown in the above, the same phonetic symbols are used for voiced and voiceless stops in the two languages, but the physical features of their segments differ from one another. Even if the two languages are said to have a voicing contrast, the stop voicing features are different from each other in a phonological framework.

That Korean L2 learners use Type 1 stops for English and Japanese voiced stops /b, d, g/ implies that they take voicing features rather than tensing ones, and the phonological features such as voicing are more relevant for the distinction than tensing. Further, as noted in Tables 2 and 3, they pronounce English and Japanese stops with a voicing delay and this means that a voicing lead is not necessarily a condition for the voiced feature and voiced sounds are pronounced with a long timing dimension from a considerable voicing lead to a voicing delay. Stop voicing is basically a laryngeal feature and should be examined in terms of timing features, which are closely related to the movements of vocal folds.

VOT is considered to be an efficient feature to specify the voiced/voiceless distinction, and voicing lead may be closely related with voicedness, while voicing delay may be related with aspiration and

voicelessness. But this specification is not appropriate, since a number of voiced categories in Japanese and English uttered by Korean L2 learners are in the range of voicing delay. It may be necessary to have revised features to account for the voiced/voiceless distinction in case of Korean L2 learners. Along with this proximity in VOT values, it can be said that Type 1 stops are used in Korean loan words from English to represent voiced stops in English, and this may be due to the phonetic similarity between Type 1 stops in Korean and voiced ones in English.

6. Summary

The present study shows that the three-way contrast of Korean stops is distinguished by a timing dimension of VOT, and Korean L2 learners use their tense unaspirated stops for the voiced category in English and Japanese, and they use strongly aspirated stops for voiceless stops in English. However, it is not clear whether they use weakly or strongly aspirated stops for the voiceless category in Japanese. Proximity in VOT values can be considered to be one of the main reasons of using a Korean stop category in learning L2 categories. Korean is characterized as an aspiration-oriented language, and their stop categories are well suited to acquire the two-categories of stop consonants in English. Three phonological features such as /VOICE, ASPIRATED, TENSE/ can be proposed and used to specify the phonetic nature of stop consonants in these three languages.

7. Acknowledgements

We would like to thank Professor Masayoshi Kakudo of Osaka University for arranging the subjects in the present study. The research was supported by the Grant-in-Aid for Scientific Research by the Japan Society of Promotion of Science Fund (22520593) in 2010.

Notes

1 There is a disagreement over labeling of the tensing feature to Type 3 stops. Kim (1965) classifies them [tense], while Lisker and Abramson (1964) classify them as [lax].

2 According to Cho (1967), Type 2 stops become voiced in most of the intervocalic positions, except a sequence such as /o _ a/ or /u _ ə/.

3 Their English proficiency was fair in speaking, and their Japanese proficiency was excellent in their own judgment.

4 See Shimizu (1996). Shimizu examines Fo and its contour, spectral analysis, and the onset frequency of the first formant as a feature to distinguish a three-way contrast of Korean stops.

5 See Shimizu (1999).

6 The VOT values associated with voiceless velar stops are always longer than those associated with bilabial and alveolar stops. There are mainly two reasons which have been often pointed out in the phonetics literature; namely

(1) the volume of the supralaryngeal cavity and (2) the motion speed of the tongue body.

References

Cho, S-B, *A Phonological Study of Korean, With a Historical Analysis*, ACTA Universitatis Uppsaliensis, Uppsala, 1967.

Dart, S. N., "An aerodynamic study of Korean stop consonants: Measurements and modeling," *J. Acoust. Soc. Am.* 81, 138–147, 1987.

Han, M. S. and Weitzman, R. S., "Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/," *Phonetica*, 22, 112–128, 1970.

Hardcastle, W. J., "Some generalizations on the tense-lax distinction in initial stops in Korean," *J. Phonetics* 1, 263–272, 1973.

Hirose, H., Lee, C. Y., and Ushijima, T., "Laryngeal control in Korean stop production," *J. Phonetics* 2, 145–152, 1974.

Iverson, G. K. and Salmons, J. C., "Aspiration and laryngeal representation in Germanic," *Phonology 12*, 369–396, 1995.

Kagaya, R., "A fiberscopic and acoustic study of the Korean stops, affricates and fricatives," *J. Phonetics*, 2, 161–180, 1974.

Kim, C-W., "On the autonomy of the tensity features in stop classification (with special reference to Korean stops)," *Word* 21, 339–359, 1965.

Shimizu, K., *A Cross-Language Study of Voicing Contrasts of Stop Consonants in Asian Languages,* Seibido Publishing Co, 1996.

Shimizu, K., "A Study on Phonetic Characteristics of Voicing of Stop Consonants in Japanese and English,"(written in Japanese) *J. of Phonetic Society of Japan*, 3(2), 4–10, 1999.

Utuki, Akira, "Plosives in Japanese and Korean: A Review of Acoustic Phonetic Studies, *Hokkaido Gengo Bunka Kenkyu*, 7, 11–27, 2009.

Zue, V., *Acoustic Characteristics of Stop Consonants*, Indiana University Linguistics Club, 1980.