A STUDY ON VOT OF INITIAL STOPS IN ENGLISH PRODUCED BY KOREAN, THAI AND CHINESE SPEAKERS AS L2 LEARNERS

Katsumasa Shimizu

Nagoya Gakuin University, Japan
kshimizu@ngu.ac.jp

ABSTRACT

The present paper examines phonetic characteristics of the initial stop consonants of English uttered by ESL learners whose native tongues are Korean, Thai, and Mandarin Chinese. It examines VOT (voice onset time) values of initial stops in the three languages and compares them with those of English stops produced by ESL learners of the three languages. From the acoustic analysis, it can be said that the voicing contrast of initial stops in Korean, Thai, and Mandarin Chinese is distinguished by a timing dimension of VOT, and that the ESL learners of these languages tend to use their stop categories to produce voiced-voiceless categories of English stops. Among three language groups, however, Korean ESL learners produce English voiceless stops with VOT values that are intermediate to those of their lax and aspirated stops. Thai and Chinese ESL learners tend to use their L1 voicing contrasts to produce the voicing contrast of English stops. Proximity in VOT values can be considered to be one of the main cues of using their L1 stop categories in learning L2 categories.

Keywords: Korean, Thai, Mandarin Chinese, VOT, ESL learners

1. INTRODUCTION

The paper examines phonetic characteristics that Korean, Thai and Mandarin Chinese ESL learners exhibit in the process of acquiring stop voicing contrast in English. These three languages are known to have different voicing contrasts from English. Korean has a three-way contrast of tense, lax, and aspirated stops, Thai also has a three-way contrast of voiced, voiceless unaspirated, and aspirated stops, and Mandarin Chinese has a two-way contrast of voiceless unaspirated and voiceless aspirated stops in their stop series.

The contrast in these languages 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. Among these languages, Korean stops have been extensively studied, while Thai and Chinese stops have been studied in details, but not so in details as done in Korean. Physiologically, Korean stops have been studied through electromyographic and fiberscopic analysis, and specific activities of laryngeal muscles have been known to classify the stop series in Korean (Dart [4], Kim [7]). Further, acoustically there have been a number of studies for stops in these languages. Han and Weitzman [6] and Cho, et al. [3] examined VOT of Korean stops. Abramson [1] and Shimizu [8] examined the acoustic characteristics of Thai stops, and Zhu [9] examined Chinese stops in details. Through these studies, phonetic characteristics of stops in these languages have been well documented, but the problem is the lack of examination of how the speakers of these languages manifest their characteristics of English stops in their learning English as a second language.

The present study was undertaken to examine the voicing contrast of the stops in these three languages and how ESL learners of these languages exhibit phonetic characteristics in learning English stops. English is quite well known to have two types of stops: voiced /b. d, g/ and voiceless /p, t, k/, and it will be significant how stop contrasts in the three languages are mapped onto a two-way contrast in English stops, since it will clarify the laryngeal relationship in these languages, and the study will reveal how the contrasts in the three languages are correlated to the stop contrasts in English in their learning as a second language.

2. EXPERIMENTAL METHOD

2.1. Subjects

The subjects in the present study are native speakers of the three language groups, and the number of subjects is as follows:

Korean	Seven (3 males, 4 females)
Thai	Four (4 females)
Chinese	Two (one male and one female)

All of them are standard speakers of these languages, and are aged 22 – 30 years, and they are studying at the universities in Japan. They studied English for almost 8 years in their own countries, and have resided in Japan for a period of 1 to 2 years. Their English proficiency is at a beginning or intermediate level.

2.2. Stimuli and procedure

The linguistic stimuli used in the present study are minimal triplets and minimal pairs embedded in the carrier sentence. The materials were written in their own characters and were presented to each subject. The stimuli were read twice by each subject. They were shown as follows:

Korean

[p*ul] horn	[pul] fire	[pʰul] grass
[t*am] sweat	[tam] wall	[tʰam] envy
[koŋ] zero	[kʰoŋ] bean	
[k*ul] honey	[kul] oyster	
[t*al] daughter	[t*am] sweat	[t*uda]float
[pada] sea		
[tal] moon	[to:l] stone	
[kat] hat	[kot] place	[kilda] be long
[pʰal] arm	[pʰi] blood	
[tʰal] a mask	[tʰək] chin	

Thai

[bà:] the shoulder	[pà:] the forest	[pʰà:] to cut
[dōn] to inspire	[tōn] the self	[tʰōn] to endure
	[kài] the chicken	[kʰài] the eggs

Chinese

[bā] eight	[pā] to lie on
[bà] father	[pà] to fear
[dā] to frame	[tā] he
[dà] big	[tà] to stamp
[gàn] to do	[kàn] to see
[gà] to hesitate	[kǎ] card

English

[pɪg] pig	[bɪg] big
[tɪk] tick	[dɪk] Dick
[kɪk] kick	[gɪl] gill

2.3. Acoustic analysis

Acoustic analysis of the recorded materials was made using AcousticCore 8 of Arcadia, Inc. The recorded audio data was digitized and analyzed at a sampling rate of 8 kHz, and they were stored for reviewing and listening. The measurement of duration such as VOT was made by manually positioning two cursors in the display of the waveform and wide-band spectrogram.

3. EXPERIMENTAL RESULTS

The measurement of VOT was made for the interval between the consonant release and the onset of voicing. It was made for three types of Korean and Thai stops and for two types of Chinese stops, and was also made for two types of English stops uttered by each language group of ESL learners.

3.1. Korean ESL learners

Table 1 indicates the mean VOT value (ms) for the three types of Korean stops uttered by Korean subjects, and Table 2 does the ones of the two categories of English stops by them.

Table 1: Mean VOT values of Korean stops (ms) (N=24) (s.d. in parenthesis).

Tense stops	Lax stops	Aspirated stops
p* 19 (7.9)	p 55 (21.8)	pʰ 82(23.0)
t* 16 (7.5)	t 51 (18.1)	tʰ 78 (26.2)
k* 33(11.3)	k 63 (20.2)	kʰ 81(14.5)

$$F(2,6)=62.89, p<.01$$

Table 2: Mean VOT values of English stops by Korean subjects (N=24) (s.d. in parenthesis).

	E-Voiced	E-Voiceless
b	24 (14.9)	p 64 (31.7)
d	18 (23.0)	t 67 (18.4)
g	38 (16.8)	k 84 (31.3)

$$t(18.48)=4.36, p<.01$$

From Table 1, it can be seen that VOT values increase in the order from tense to aspirated stops, and there is a difference among the three categories of Korean stops. It is apparent that aspirated stops are produced with considerable delay of voicing and are strongly aspirated. As the standard deviation shows, however, there is an overlap between lax and aspirated stops. Table 2 shows that Korean ESL learners clearly distinguish the two categories of English on the VOT dimension. It can be said that voiced stops in English are produced with some voicing delay, and they are produced with VOT values which are close to those of tense stops in Korean. Further, it can be pointed out that voiced stops in English are articulated with positive VOT values, and that Korean ESL learners use their tense stops to produce English voiced stops, while they produce English voiceless stops with VOT values that are intermediate to those of their lax and aspirated stops.

3.2. Thai ESL learners

Table 3 indicates the mean VOT value (ms) for the three types of Thai stops, and Table 4 does the one for the two types of English stops uttered by Thai subjects who are learning English as a second language.

Table 3: Mean VOT values of Thai stops (ms) (N=24) (s.d. in parenthesis).

Voiced	vl.unaspirated	vl.aspirated
b -87(16.9)	p 11 (6.1)	p^h 87(20.8)
d -75(3.7)	t 10 (3.7)	t^h 91(27.6)
	k 21 (20.0)	k^h 113(19.5)

Table 4: Mean VOT values of English stops by Thai subjects (ms) (N=12) (s.d. in parenthesis).

E-Voiced	E-Voiceless
b -100 (23.7)	p 74 (23.0)
d - 68(57.6)	t 91 (31.1)
g - 25 (24.9)	k 94 (21.8)

From Table 3, it can be seen that three types of stops in Thai can be clearly distinguished by VOT mean values. The grouped mean values show a tri-modal distribution of VOT along a single timeline for three types of voicing categories. Voiced stops are produced with a prevoicing, voiceless unaspirated stops are produced with a short delay of voicing, and voiceless aspirated stops are produced with a delay of voicing. Further, as is known in other languages, there is a tendency that the value is greater for velar stops than for bilabial and alveolar stops.

Table 4 shows the mean VOT values of English stops produced by Thai ESL learners, and English voiced stops are produced by prevoicing, while voiceless stops are produced with a considerable delay of voicing. It can be said that voiced stops in English are produced with VOT values which are close to those in Thai, while voiceless stops are produced with VOT values which are close to those of aspirated stops in Thai.

3.3 Mandarin Chinese ESL Learners

Table 5 indicates the mean VOT values of Chinese stops and Table 6 does the one of English stops produced by Chinese ESL learners.

As is known, Mandarin Chinese has two types of stops: voiceless unaspirated and voiceless aspirated stops. Table 5 clearly shows the difference in VOT values for two types of stops, and there is no overlapping in VOT between the two types of stiops. Aspirated stops are

considerably delayed in the timing of voicing, and the results show the general characteristics of the two types of stops. From Table 6, it can be seen that Chinese ESL learners produce English voiceless stops with the VOT values close to their voiceless aspirated stops, while they do English voiced stops with the VOT values which are close to their unaspirated stops.

Table 5: Mean VOT values of Chinese Stops (N=12) (s.d. in parenthesis).

Unaspirated	Aspirated
p 11 (4.7)	p^h 83 (15.4)
t 17 (8.8)	t^h 88 (20.0)
k 19 (3.6)	k^h 80 (16.5)

Table 6: Mean VOT values of English Stops by Chinese ESL Learners (N=12) (s.d. in parenthesis).

E-Voiced	E-Voiceless
b 16 (6.4)	p 76 (29.5)
d 24 (23.1)	t 65 (13.1)
g 25 (11.1)	k 93 (15.5)

4. DISCUSSION

The examination of VOT reveals that VOT values are functional in distinguishing the stop contrasts in Korean, Mandarin Chinese, and Thai. In Korean, a three-way stop can be distinguished by the VOT dimension: VOT in word-initial position is short for tense stops, intermediate for lax stops and long for the aspirated stops. However, the value ranges overlap in two out of seven subjects, and these results accord with previous studies (Cho, et al.[3]). Other acoustic parameters such as Fo and Fo contour and voice quality of the following vowels may be relevant for classifying the contrasts. In examining VOT values of their English stops, Korean ESL learners closely align their tense stops with English voiced stops, and they try to merge their lax and aspirated stops to produce voiceless stops in English.

As to Thai stop contrast, it can be said that there is a clear-cut difference for three types of voicing categories in VOT values. The range of voiced categories is rather extensive as shown in standard deviation, but there is no overlapping with other categories. Further, Thai ESL learners produce English voiced stops with the VOT values which are close to their voiced stops and do English voiceless stops with the ones which are close to their aspirated stops.

Further, in Mandarin Chinese, VOT is functional in distinguishing the two stop categories

which are both voiceless. Voiceless unaspirated stops are produced with a short VOT value, while voiceless aspirated stops are produced with a considerable delay of voicing. They use their voiced stops to align with English voiced ones, and use their aspirated stops to align with English voiceless ones.

In these three languages, VOT is a major cue to distinguish the voicing categories, and in acquiring English as a second language, the subjects in this study use VOT as a main feature in acquiring the categories of L2 stop consonants. Proximity in VOT values can be considered to be the main factor of using an L1 stop categories in leaning L2 ones.

In learning L2 speech sounds, there are several models such as speech learning model and perceptual assimilation model (Best, et al. [2], Flege [5]. In these models, phonetic distance of stop consonants in the three languages with English ones is the major cue in learning new L2 sounds. The subjects in these language try to find L1 sounds whose VOT values are close to L2 ones, and they use their own stops in pronouncing English voicing categories.

5. SUMMARY

The present study shows that the stop voicing contrast of initial stops in Korean, Thai, and Mandarin Chinese is distinguished by a timing dimension of VOT, and the ESL learners in these languages use their stop categories in producing two categories of English stops whose VOT values are close to their own stops. Proximity in VOT values can be considered to be one of the main cues of using their L1 stop category in learning L2 categories. Korean, Thai and Mandarin Chinese stops are characterized as an aspiration-based category, and their categories are well suited to acquire the two categories of stops in English.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Abramson, A.S. 1972. Word-final stops in Thai. In Harris, J.G., Noss, R.B. (eds.), *Tai Phonetics and Phonology.* Central Institute of English Education, Office of State Universities, Bangkok.

[2] Best, C.T., McRoberts, G.W., Goodell, E. 2001. American listeners' perception of nonnative consonant contrasts varying in perceptual assimilation to English phonology. *J. Acoust. Soc. Am.* 109, 775-794.

[3] Cho, T., Jun, S.A., Ladefoged, P. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *J. Phonetics* 30, 193-202.

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

[5] Flege, J.E. 1987. The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *J. Phonetics* 15, 47-65.

[6] Han, M.S., Weitman, R.S. 1970. Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica* 22, 112-128.

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

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

[9] Zhu, C. 2010. *Chuugokugo Nihongo no Jikkenteki Kenkyuu (An Experimental Study of Chinese and Japanese Speech Sounds).* Kuroshio Publishing Co.