東義大學校東義論集
第20輯 人文·社會科學編
1993. 2. pp. 41~59

A Voice Onset Time Comparison of English and Korean Stop Consonants

Byunggon Yang*

Contents

1. Introduction

The stop consonants in Korean may be classified into three groups according to the place of articulation. They are labials, dentals and palatals. For each group, there are three different types according to the manner of articulation. They are forced, aspirated, and unaspirated stops. Similarly, English has three groups of stops according to the place of articulation. It has two different types according to the manner of voiced/voiceless distinction. Therefore, it would be very interesting to study how Korean students of English produce those Korean and English stops cross-linguistically.

The aim of this study is to compare Voice Onset Time (Hereafter simply VOT) values of English and Korean stops. VOT values of English stops at the initial

* 인문대학 영어영문학과 전임강사

position produced by Korean students and American students attending the University of Texas at Austin were compared to find out statistically significant difference in their productions. Also the Korean stops at the initial positions were measured to check any phonetic interference in the production of English stops.

2. Review of the Literature

2.1 Voice Onset Time

VOT is defined as "the time with respect to release for the onset of voicing" (MacKay 1987 : 93-4). Lisker and Abramson(1964 : 389) adopted the convention of assigning zero-time to their reference point : the instant of release. Measurements of VOT before the release are stated as negative numbers and called "voicing lead," while measurements of VOT after the release are stated as positive numbers and called "voicing lag"(MacKay, 1987 : 94). If release and voicing are simultaneous, VOT is zero. Fig. 1 shows its relationship.(The point X is plosive release.)

In a cross language study Lisker & Abramson(1964 : 388-39) examined how well VOT serves to separate the stop categories of eleven languages. They limited their study to word-initial position before vowels. From a comparison of the mean values and ranges of stops of eleven languages they found that VOT was a reliable parameter for differentiating the voicing contrasts in several languages. Also, Lisker & Abramson(1971 : 770) stated that VOT is "the single most effective measure for classifying stops into different phonetic categories with respect to voicing". There are three relatively specific ranges of VOT for initial stop

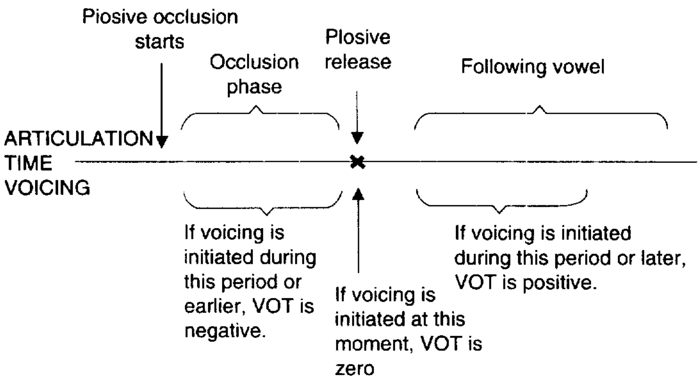

Fig. 1. Voice Onset Time.

consonant productions. In the voicing lead range, voicing precedes the release of stop and these stops have a negative VOT value. Voicing begins at 0 to +20 ms after the stop release in the short voicing lag range and +40 ms or more after the release for the long voicing lag range(Kewley-Port & Preston, 1974 ; Lisker & Abramson 1964, 1971).

Keating(1984 : 295) also classifies long and short lag as follows : Positive VOT values to about 20-35 msec(depending onthe place of articulation) are called 'short lag' ; higher values arecalled 'long lag'. In general, stops traditionally described as being voiceless unaspirated(or devoiced) have short lag VOT' s, while stops traditionally described as being voiceless aspirated have long-lag VOT' s.

MacKay(1987 : 95) shows VOT values in English word-initial voiced plosives. The values are usually around zero, either slightly negative or moderately positive as shown in Fig. 2.

A brace in Fig. 2 indicates the possible range where voicing might start ; once started, it continues throughout the following vowel. The value is often positive, which means voicing does not start until after the release. The occlusion phase is thus voiceless, although this is a "voiced plosive" (MacKay, 1987 : 95).

Lisker & Abramson(1967) reported that vowel identity had no systematic influence on VOT. However, Klatt(1975 : 691) observed a "statistically significant" difference in VOT values for voiceless stops, depending on whether high- or mid-vowel followed the stop. Klatt suggested that the failure of the Lisker & Abramson' s study to detect this vowel effect may be related to the limited scope of the former authors' data.

Klee, Weismer & Ingrisano(1976) reported on a limited set of data in which VOT associated with the first consonant in a CVC word was altered systematically by the voicing characteristic of the final consonant. In addition,

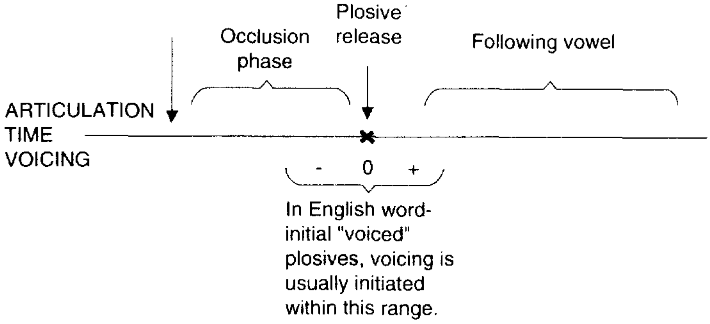

Fig. 2. English VOT range.

the Klee et al. results appeared to be in conflict with the trend of vowel conditioning on VOT found by Klatt, in that a low vowel /æ/yielded longer VOT's than a high vowel /i/. Weismer(1979 : 198) explained the discrepancy above by the types of vowel used in the two studies :

> Klatt's tongue-height contrast involved two-high, tense vowels /i/, /u/ and tense /ay/ and lax /e/ mid-vowels, whereas the Klee et al. contrast was based on high /i/ and low /æ/lax vowels. Thus, it could be reasoned that an unidentified influence on VOT of the tense/lax distinction in vowels might produce the inconsistency in the findings of the two studies.

Weismer's(1979 : 202) study has also shown that "when other factors are held constant, VOT of a word-initial stop is related systematically to features of the following vowel and final consonant in the word. Specifically, VOT will be longer when the vowel is tense, as compared to lax, and when the final consonant is voiced, as compared to voiceless."

On the other hand, voiceless plosives in the word-initial position will be aspirated and have a greater positive VOT. If a voiced plosive is word-medial between two voiced segments, any measure of VOT for this plosive is meaningless since the onset of voicing occurred with a previous segment. If a voiceless plosive is situated intervocalically at the beginning of a stressed syllable, it is aspirated and has a relatively long positive VOT. If it is situated at the end of a stressed syllable or at the beginning of an unstressed syllable it is generally not aspirated and has rather a shorter VOT, more like that of a voiced plosive. Voiceless plosives following an /s/ in word-initial position, are not aspirated and have shorter VOT values, falling into the range of voiced plosives in word-initial position so they are heard as voiced plosives if put in word-initial position. Various cases of VOT in different environment are summarized in Table 1-1 (Mackay, 1987 : 96).

2.2 English stops

English is said to have six plosive phonemes : voiceless and voiced bilabial /p, b/, alveolar /t, d/, and velar /k, g/. The term stop is also used for plosives. MacKay (1987 : 90) explains the mechanism of English plosives :

> A plosive consonant is formed by blocking the oral cavity at some point.

Table 1-1. Voice Onset Time (VOT) of voiced and voiceless English in various environments

Position	Voiceless	Voiced
Word-initial Examples	pan tan can	ball doll Gaul
Characteristics	Aspirated VOT : long, positive	Not aspirated VOT : slinghtly negative through 0 to moderately positive
After word-initial/s/ Examples	spare stick skate	N. A.
Characteristics	Not aspirated VOT : 0 to moderately positive	
Intervocalic, not at the beginning of stressed syllable Examples	kaput retake(vb) recoil(vb)	rebuff redo(vb) regard
Characteristics	Aspirated VOT : long, positive	Not aspirated VOT : not measurable ; continuous voicing through preceding and following vowels
Intervocatlic, not at the beginning of stressed syllable Examples	capping mater makeup	rubber madder Maggy
Characteristics	Not aspirated VOT : 0 to somewhat positive. /t/ may be flapped : VOT then unmeasurable.	Not aspirated VOT : not measurable ; continuous voicing. /d/may be flapped.

During the articulation of most plosives the velum is raised, blocking off the nasal passages. This permits a certain amount of pressure to build up in the oral cavity behind the occlusion……. The pressure that is built up is released suddenly, as a minor "explosion" or "popping". The plosives cannot be pronounced alone.

MacKay(1987 : 91) lists the three usual phases of plosive articulation :

1. The shutting or closing phase, during which the articulators are moving from a previous open state to the closed state.
2. The closure phase or occlusion, the momentary total blockage of the vocal tract.
3. The release, which is the abrupt reopening of the vocal tract.
   It is sometimes accompanied by a burst of air, if the intraoral pressure has been great.

The plosives /p/ and /b/ are bilabial sounds : the blockage of the oral cavity is made with the two lips. /p/ is a voiceless sound : the vocal folds are not vibrating. /b/ will be pronounced with a similar articulatory gesture with glottal vibration. For the plosives /t/ and /d/, the oral cavity is blocked by placing the blade of the tongue against the alveolar ridge. The sides of the tongue are placed along the upper teeth, completing the seal. The release occurs when the blade of the tongue is lowered, allowing the pressure to explode over the top. In the articulation of the two velars /k/ and /g/, the tongue dorsum is raised to meet the velum ; the velum is not lowered to meet the tongue. The nasal passages must remain closed.

2.3 Korean Stops

Korean stops are classified into three categories while English stops are classified into two categories(Lisker and Abramson, 1964 : 397). /p,t,k/ have been called voiceless, tense, long, and glottalized. The series /P, T, K/ are said to be in initial position, voiceless, lax and slightly aspirated. /p^h,t^h,k^h/are described as voiceless and heavily aspirated, but lax.

Han and Weitzman(1970 : 112) classify Korean stops as follows :

> In Korean, a three-way distinction in both manner and place of articulation serve to differentiate nine stop consonant phonemes : the aspirated bilabial, dental, and velar stops, /p^h, t^h, k^h/ ; the weak bilabial, dental, and velar stops, /p, t, k/ ; and the strong bilabial, dental, and velar stops, /P,T,K/. All Korean stops may occur in initial position ; and in this position, they are all voiceless.

According to their study on three native speakers of the Seoul dialect of

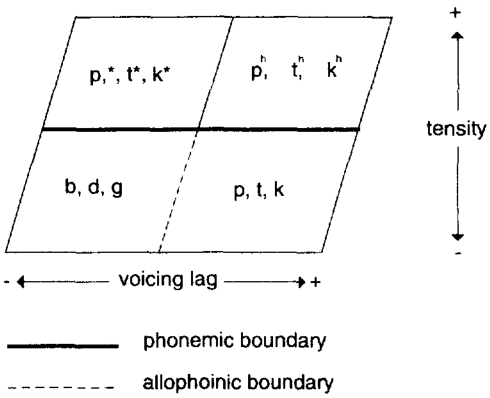

Fig. 3 Three Series of Korean Stops.

Korean, VOT averages two to five times longer for aspirated stops than for weak stops. The VOT of weak stops averages one and a half to five times longer than those of strong stops.

Kim(1965 : 356) arranges the three series of Korean stops in Fig. 3.

In this diagram /p, t, k/ can move freely to and fro along the voicing scale, but they cannot move up across the tensity boundary line. Kim(1965 : 356) concluded : "In Korean stops, tensity is the primary feature, the feature that occupies the higher node in the phonological hierarchy of stops, and voice onset timing the secondary." Similarly, Kagaya(1974 : 161-65) classified the three types of Korean stops as "forced", "lax" and "aspirated". He studied two Korean speakers' productions in a /CV/ environment and found that :

1. In the aspirated type, an increase in glottal width is observed in the initial phase of utterance and the glottis closes rapidly for the following voiced segment. The articulatory explosion occurs around the moment when the glottal width reaches its maximum value.
2. In the lax type, the glottis begins to close gradually with some characteristic fluctuations in glottal width after the articulatory explosion, a rapid reduction takes place for the following voiced segment.
3. In the forced type, glottal width seems always to decrease monotonically in this phonetic environment. Complete contact of the vocal processes is achieved at a point from 80 to 100 ms before the articulatory explosion. During this period,······ a spindle-shaped gap is always observed in the membranous portion of the glottis.

Hardcastle(1973 : 265) found that the Korean VOT values for /p,t,k/ are "usually about 3 to 5 times longer than those for /P,T,K/. The heavily aspirated

/pʰ, tʰ, kʰ/ average" about 2 to 3.5 times longer than the /p, t, k/". Two well-known analyses establish only the second type as unit phonemes, while those of the first and third types are each represented as sequences of the two phonemes, the first type is taken either as a case of gemination or as a sequence of a simple stop phoneme and a phoneme of "glottal tension", while the third type is a sequence of a simple stop followed by the phoneme /h/.

2.4 Cross-Linguistic Studies on VOT

It is commonly accepted that the second language(L2) learners "identify" L2 phones in terms of native language(L1) categories and, as a result, use articulatory patterns established during L1 acquisition to realize those L2 phones. James(1985) noted that nearly all L2 speech errors involve phonemes which either do not occur, or are realized differently in L2 and L1. This suggests that L2 learners may have difficulty establishing the articulatory patterns needed to produce both new and similar L2 phones authentically ; or contrary to the assumption stated above, they may continue to identify both new and similar L2 phones in terms of an L1 category. Flege & Hillenbrand(1987) hypothesized that an upper limit existed on the extent to which L2 learners approximated L2 phonetic norms for similar phones because L2 learners "merged" the phonetic properties of similar L1 and L2 phones within a single category. L2 learners have been observed to produce /t/ in L2 with "compromise" VOT values that were longer than the values typical for L1, but shorter than the values typical for L2 (Caramazza, 1973 ; Williams, 1980). Flege & Hillenbrand(1984) found that none of the English speakers producing French were able to produce authentically the short-lag VOT values characteristic of French /t/. Flege(1986) found that even the experienced speakers produced VOT values for French which were significantly longer than those for monolingual French speakers.

Flege(1987 : 187) hypothesized that adult L2 learners' failure to make effective use of the sensory information associated with certain phones in L2 is the result of a mechanism termed "equivalence classification". If short-lag and long-lag realizations of /p, t, k/ in L1 and L2, respectively, are judged to be realizations of the same categories, L2 learners will not establish new phonetic categories for the long-lag L2 stops, even though the acoustic differences distinguishing the corresponding L1 and L2 stops may be detected auditorily. This mechanism of equivalence classification leads them to identify acoustically different phones in L1 and L2 as belonging to the same category. This may prevent them from producing similar but now new phones authentically. This hypothesis may also lead to a strong claim that adult L2 learners will never match native English speakers in producing stops. Previous research indicated that groups of subjects who learned English after about the age of 5-6 years showed phoneme boundaries at VOT values that were often intermediate in value to the norms established for the L1 and L2(Flege et al., 1987 ; Williams, 1979). In some instances individual subjects have category boundaries at values that are genuinely "intermediate" to the L1 and L2 norms, suggesting the phonetic merger(or assimilation of the properties of corresponding stops in L1 and L2. Non-native adults with relatively little English language experience produce the long-lag /p,t,k/ of English with short-lag VOT values if /p,t,k,/ are realized as voiceless unaspirated stops in their L1(i.e., as short-lag stops). Those with relatively more L2 experience typically produce English /p,t,k/ with VOT values intermediate to the long-lag VOT norm or corresponding stops in their L1(Flege,1987 ; Major, 1987 ; Port et al., 1980). Flege & Eefting(1987 : 78) also found that the mean VOT for /p,t,k/ observed for the native Spanish subjects in a group of later childhood bilinguals was almost "perfectly intermediate(57 ms) to the means observed for the Spanish adults who did not speak English(26 ms) and the monolingual English adults(87 ms)."

Williams(1979) found that 8-10 and 14-16 year-old native Spanish(Puerto Rican)children produced English /p/ with Spanish-like mean VOT values of about 40 and 20 ms, respectively. It seems that adults who learn English L2 as young children, but not those who learn English as an L2 later in life, will succeed in realizing English /p,t,k/ authentically. Early L2 learners would have the potential advantage of being exposed to tokens of English[p^h, t^h, k^h] at a time when it may be relatively easy to establish new phonetic categories, and when they have many years to refine the phonetic realization rules needed to output those phonetic categories. Two recent studies support the early L2 learning hypothesis. Mack(1983) did not observe a significant VOT difference for /t/ between English monolinguals and adults who had learned French and English as young children ; moreover, Williams(1977) found no VOT difference for /p/ between English monolinguals and adult subjects who had learned Spanish and

English as young children. Caramazza et al.(1973), on the other hand, found that French Canadians who began learning English in school "no later" than the age of seven years produced /p, t, k/ in English with VOT values that were significantly shorter(51 ms) than English monolinguals(74 ms).

Major(1987 : 197) claimed that the fine phonetic differences between the two languages can cause the learner considerable difficulty. One such difficulty is learning the timing relations between glottal and supraglottal articulations or VOT. Major(1987 : 201) demonstrated in his study that "global[overall] foreign accent is significantly correlated with VOT : as accent becomes more native like VOT proficiency" and suggested that "at least this aspect of second language acquisition is within the grasp of learners."

3. Method

3. 1 Subjects

A total of 14 subjects between the ages of 20 to 31 years were selected to participate in the recording session at the University of Texas at Austin(UT). They were grouped into two : seven American students majoring in broadcasting and journalism and seven Korean students majoring in linguistics. The recording by American students was made in a sound-proof booth in the Communication Building. The Korean subjects' recording was made in a quiet room in the UT Phonetics Lab.

3. 2 Speech Samples

The speech samples consisted of 30 English words and 24 Korean words. English samples were mono- or di-syllabic words initiated by /b, d, g, p, t, k/. The immediately following environments were : Stops in a stressed syllable : before vowel /a, i, a : / ; after a consonant *s* ; Stops in an unstressed syllable : before vowel /i, ə / ; after a syllable *re*. The sample words were arranged in the order of /b, d, g, p, t, k/ in each different environments on a page. The vowel [a] and [i] are selected because they have opposite features such as [+low, +back], [+high, -back], respectively. Also the following consonant after the vowel was set to [t] or [P] to assimilate Korean counterparts. This was done to obtain as similar word samples as possible to Korean ones. Korean samples were di- or tri-syllabic words initiated by /p, t, k, P, T, K, p^h, t^h, k^h/. Three different environment were chosen : before vowel /a, i, a : /.

3.3 Procedure

American and Korean Subjects were asked to produce English stop consonants. In addition to English words, Korean subjects produced Korean counterparts. Before the recording session they were given the following preliminary instructions : "The material you are recording contains English words(and Korean words for Korean subjects). Please produce them as casually but clearly as possible. Please read each word twice at level tone, and pause a second after each word." Korean subjects produced English sample words before the Korean equivalents in order to reduce the direct interference from the Korean. After finishing the last part of the English samples, subjects were asked to put a short half-minute pause and continue reading the following Korean words. The experimenter demonstrated the pronunciations of each word and practiced how to start and stop. Then, the experimenter cued each subject to start recording. It was necessary to have subjects read the words in isolation because a Korean phonological rule causes /b, d, g/ to be realized as approximants in many contexts other than absolute utterance-initial position(Harris, 1969). The recordings by American and Korean students lasted about 2 minutes for the Americans and 4 minutes for the Koreans.

3.4 Acoustic Measurements.

The audio signal was reproduced using a Superscope CD-330 professional cassette recorder, amplified by a Denon DMA-730, and passed through a White Instruments Model 4382 anti-aliasing filter(8 kHz cutoff), before being digitized at a 20 kHz sampling rate via the DEC MNCAD 12-bit analog-to-digital converter, and stored on disk. The computer used for all input and analysis was a Digital Electronics Corporation MINC-11/23.

A Tektronix 4006-1 graphics terminal(Storage screen type) was employed for high-resolution graphic display, a Tektronix 5103N Oscilloscope, driven by two D/A channels, is used for animated computer display of wave forms(refreshed 60 times per second), and a Tektronix 4632 Video Hardcopy Unit copies from the VT 105 screen.

A family of programs, written in U.T. Phonetics Lab and known collectively as SCOPE, was used to interactively examine, measure and analyze the signals. A joy stick was used to control velocity and direction of movement through the audio file. Through vertical movement of the joy stick it was possible to 'zoom in', to examine a small stretch of sound, or 'zoom out' for a more global per spective. Audio monitoring of the signal was made to check the sound of each segment. VOT determinations were made with a precision that permitted rounding to the nearest 0.5 ms. VOT was measured to the nearest 0.1 ms from the display of a high graphics terminal. The left cursor was set at the sharp increase in wave form energy which signaled the release of /b, d, g, p, t, k/. The right cursor was set at the first upward-going zero crossing which signaled voicing onset. If the subject produced VOT lead, the right cursor was set at the first burst of /b, d, g/.

4. Results and Discussion

4.1. English VOT Values

The numerical data for the initial stops of isolated words for English and Korean were collected and listed as follows : The first line gives average values of VOT for each category ; the second line shows the ranges of values observed ; and the third line indicates the number of tokens of each stop phoneme. Positive and negative values are given for /b,d,g/ by a slash. Lisker and Abramson (1964 : 395) give the reason as follows :

> To have given a simple set of values would have meant lumping positive and negative values of VOT as items of a single population, and it appeared rather that these are distributed within two discontinuous ranges. In such a situation it would be misleading to determine single average values of voice onset time for the members of the /b,d,g/ set.

The following average VOT values were calculated from the pooled data. The total tokens for English /b,d,g/ were 28 in the environment of vowels /a,i,a : /, unstressed initial syllable. As was described before, the VOT measurement of /b, d,g/ between the voicing segment is meaningless. Also there is no word which begins with sb-, sd-, and sg-. For /p,t,k/, the mean value is obtained from the environments of vowel /a, i, a : /, unstressed initial syllable and stressed

Table 1. VOT in Msec : English produced by seven American students.

	[b]	[d]	[g]
AV.	17/-78	20/-91	32/-78
R.	9 : 32/-173 : -12	14 : 25/-210 : -26	15 : 57/-148 : -32
N.	3/25	3/25	6/22

*AV : average values of VOT for each category. R : Range N : Number of tokens.

Table 2. VOT in Msec : English produced by seven American students.

	[p]	[t]	[k]
Av.	77	95	88
R.	47 : 142	54 : 193	45 : 131
N.	35	35	35

Table 3. VOT in Msec : English produced by seven Korean students.

	[b]	[d]	[g]
Av.	15/-104	19/-13	31/-16
R.	9 : 31/-150 : -16	9 : 39/-13 : 0	17 : 43/-16 : 0
N.	24/4	27/1	27/1

Table 4. VOT in Msec : English produced by seven Korean students.

	[p]	[t]	[k]
Av.	63	70	84
R.	33 : 114	36 : 107	41 : 146
N.	35	35	35

Table 5. VOT Msec : English after s- produced by seven American students.

	[p]	[t]	[k]
Av.	16	25	29
R.	12 : 22	19 : 32	18 : 37
N.	7	7	7

Table 6. VOT in Msec : English after s- produced by seven Korean students.

	[p]	[t]	[k]
Av.	22	27	34
R.	7 : 45	17 : 45	14 : 62
N.	7	7	7

secondary syllable. The /p, t, k/ after /s/ is calculated independently from the above pooled data because the average values for the three stops appear quite different from the others.

4.2 Korean VOT Values

VOT was measured for Korean word samples. Korean stops at the initial position were all voiceless. The following data were obtained from pooling the data together in the environments of vowels /a, i, a : / for /p, t, k, p^h, t^h, k^h/(21 tokens each). For the stops /P, T, K/ the tokens were limited to the environment of vowel /a, i/(14 tokens each). The following table shows each average VOT range and number of tokens.

In the Korean data, VOT values for aspirated stops were five to six times those of unaspirated stops. VOT values for aspirated stops show three times those of unaspirated stops. Ranges overlap slightly but still one finds that the subjects produced the three categories of stops distinctively which is seen in the range of values made by the Korean subjects.

Table 7. VOT in Msec : Korean produced by seven Korean students.

	[p]	[t]	[k]
Av.	35	39	53
R.	14 : 65	15 : 81	31 : 66
N.	21	21	21

Table 8. VOT Msec : Korean produced by seven Korean students.

	[p^h]	[t^h]	[k^h]
Av.	75	75	95
R.	39 : 124	41 : 113	64 : 182
N.	21	21	21

Table 9. VOT in Msec : Korean produced by seven Korean students.

	[p]	[t]	[k]
Av.	12	16	24
R.	7 : 17	8 : 29	7 : 40
N.	14	14	14

4.3 VOT Comparison

In order to see the difference in the productions of the two groups, a one-way ANOVA in the StatView 512+™ was used. The following Table 10 shows each probability level rounded to the nearest 0.01 according to the different environments following the stops. The asterisk after the p-level means a significant difference at 95%.

The following Table 11 lists the comparison between English and Korean VOT

Table 10. VOT Comparison of American Students to Korean Students

Factor A*B	b	d	g	p	t	k
ESA*KSA	.01*	.00*	.00*	.04*	.00*	.05*
ELA*KSI	.02*	.00*	.02*	.19	.01*	.07
ELA*KLA	.16	.01*	.00*	.53	.14	.77
EUNST*KUNST	.00*	.00*	.01*	.84	.80	.54
ESEC*KSEC	n.a.	n.a.	n.a.	.06	.06	.82
ES-*KS-	n.a.	n.a.	n.a.	.19	.57	.46
E*K(Total)	.00*	.00*	.00*	.01*	.00*	.43

(ESA : English VOT values before the short vowel/a/. ELA : English VOT values before the long vowel /a/. EUNST : English VOT values in an unstressed syllable. ESEC : English VOT values in a secondary syllable. ES- : English VOT values after the consonant/s/)

Table 11. VOT comparison of American students to Korean students

Factor A*B	b	d	g	p	t	k
ESA*KSA	.00*	.00*	.00*	.03*	.00*	.04*
ELA*KSI	.00*	.00*	.00*	.17*	.01*	.50*
ELA*KLA	.00*	.00*	.00*	.36*	.37*	.64*
E*K(Total)	.00*	.00*	.00*	.49*	.00*	.55*

values.

The comparison of American students to Korean students in the final row of Table 10 shows a significant difference in /b, d, g, t/.($p<0.00$). The two VOT values of /p, k/ do not show much difference. This can be expected on the following table 11. The pooled comparison of languages, English to Korean, also shows a significant difference in VOT values for /b,d,g,t/.($p<0.00$) The main reason that there is so vast a difference in the /b, d, g/ would be that Korean students produced English stops with the norms of their native language. Korean lacks the voicing distinction in the initial syllable position. In fact, only one out of seven Korean subjects produced the voicing lead(-150). The mean value for Korean students was 7 for voicing lag(6 tokens) and -150(1 token) for voicing lead. Thus, the /b/ before the long vowel [a] has a p-level of .16. On the other hand, only one out of seven American students produced /b, d, g/ with a short voicing lag(+9 and+11 for /b/ ; +27 for /g/). For /t/, the pooled language comparison in Table X shows a significance level of $p<.00$. This results mainly from the VOT values before short vowels /a,i/. However, one finds also a nonsignificant difference in other environments. Korean [t] tends to be produced in the apico-dental region while English [t] tends to be produced in the alveolar region. Thus, VOT values for Korean dental [t] are shorter than those for English alveolar [t]. A future curricular plan should include the voicing feature for /b,d,g/ and correct tongue placement for the sound [t] to approximate English equivalents as a major part of English education.

Korean students produced /p, t, k/ with a significant difference from American students in the environment of English short vowel /a/ ($p<.05$).

Again, this statistical difference may be predicted when we look at the first two rows of Table 10 and 11. In other words, English and Korean subjects produced English VOT values differently because VOT values for English and Korean show a significant difference. On the other hand, there is no significant difference in /p, k/($p>.14$) in the pooled comparison. However, the pooled comparison of American students to Koreans in Table 10 shows that in the production of /p/ there is a significant difference ($p<.01$). It seems that a few extreme values on /p/ especially before the long vowel /a/ led to this conflicting result. It must be pointed out that in Korean the role of the length of the vowel has become small as a distinctive feature. On the other hand, the difference in

American and Korean subjects' /p/ and /k/ before the short vowel /i/ was not significant($p>.07$) which was also true of the language comparison in Table 11 ($p>.17$). These figures show that Korean students' English production can be traced back to their native language.

Since stress is not a distinctive feature in Korean, the author found no Korean samples which are equivalent to English ones. This may result in a rather loose comparison between the two data. /p, t, k/ in the unstressed position showed no significant difference between the two groups of students. Korean /p^h, t^h, k^h/ are quite similar to those of English in that environment. Korean students produced English /p, t, k/ correctly in the stressed second syllable or after consonant /s/. VOT values of Korean /p^h, t^h, k^h/ are quite similar to those of English in the unstressed position and stressed second syllable. Furthermore, there is not much difference in the production of /p, t, k/ after a consonant /s/. This can be possible because Korean has stops /P, T, K/, which have almost similar VOT values to those of English(Compare Tables 5, 6, and 9). Here again, Korean subjects may have used the quality of Korean /P, T, K/ in the production of English /p, t, k/ after a consonant /s/. However, one can still find that some subjects produced /P, T, K/ with some aspiration looking at the range presented in Table 8. Korean students are likely to choose one of the three categories to approximate the two categories of English voiceless stops.

VOT shows a great deal of variance according to the following vowel. Therefore, if any two languages are compared on the pooled data, the result may be ambiguous or distorted. In this sense, VOT comparison should be made in a controlled environment such as a vowel and an immediately following consonant. Moreover, experiments with a small number of subjects may lead to an incorrect conclusion. Especially, when VOT values of the voiced segments of /b, d, g/ are compared, the original VOT values indicate too much variance in themselves. Therefore the statistical result will represent a somewhat biased one. Also the extreme score of the individual should be double checked to see whether the score is an over-articulated one. It would be better to have a proper number of subjects that will offset statistically any extreme values. The prevoicing segments or consonant segment should be compared separately to have more reliable results of /b, d, g/.

In conclusion, this study found that Korean students produced English stops with significantly different VOT values from those of American students when the Korean stops had significantly different VOT values from those of English. Furthermore, when the Korean stops had no significantly different VOT values to those of English, Korean students produced English stops with no significantly different VOT values to those of American students. In addition, we found that there exists a vast difference in the production of voiced stops. So the educational material should stress the difference in VOT values for voiced/voiceless stops. They would be ready to produce correct /P,T,K/ after the consonant /s/ in English if they can properly use the norms of unaspirated Korean stops. A cross-linguistic comparison should be carefully done since VOT values vary greatly according to the preceding and following environments.

Bibliography

Abramson, A.S. and Lisker, L.1973. "Voice-timing perception in Spanish word-initial stops." *Journal of Phonetics* 1, pp.1-7.

Barry, W.J. 1974. "Language background and the perception of foreign accent." *Journal of Phonetics* 2, pp. 65-89.

Caramazza, A., G. Yeni-Komishian, E. Zurif, and E. Carbone. 1973. "The acquisition of a new phonological contrast : the case of stop consonants." *Journal of Acoustical Society of America* 54, pp.421-28.

Kim,C.W. 1968. "The vowel system of Korean." *Language* Vol. 44. No. 3 , pp. 516-527.

_________. 1965. "On the autonomy of the tensity feature in stop classification." *Word* 21, pp. 339-359.

_________. 1970. "A theory of aspiration." *Phonetica* 21, pp. 107- 116.

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, pp.47-65.

Flege, J.E. and W.S. Brown. 1982. "The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance." *Journal of Phonetics* 10, pp.335-345.

Flege, J.E. and W. Eefting. 1987. "Production and perception of English stops by native Spanish speakers." *Journal of Phonetics* 15, pp.67-83.

_________. 1988. "Imitation of a VOT continuum by native speakers of English and Spanish : evidence for phonetic category formation." *Journal of Acoustical Society of America* 83 (2), pp.729-740.

Flege, J.E. and J. Hillenbrand. 1987. "A differential effect of release bursts on the stop voicing judgments of native French and English listeners." *Journal of Phonetics* 15, pp. 203-208.

Haag, W. K. 1979. "An articulatory experiment on voice onset time in German stop consonants." *Phonetica* 36, pp.169-181.

Halle, M. and G.N. Clements. 1983. *Problem Book in Phonology*. Cambridge, Mass : MIT Press.

Han, M.S. and R.S.Weitzman. 1970. "Acoustic features of Korean /P,T,K/, /p,t,k/ and /pʰ,tʰ,kʰ/." *Phonetica* 22, pp.112-128.

Hardcastle, W.J. 1973. "Some observations on the tense-lax distinction in initial stops in Korean." *Journal of Phonetics* 1, pp.263-272.

Hirose, H., C. Y. Lee and T. Ushijima. 1974. "Laryngeal control in Korean stop production." *Journal of Phonetics* 2, pp.145-152.

James, C. 1985. *Contrastive Analysis*. Singapore : Longman Publishers.

Jamieson, D.G. and M.F.Cheesman. 1987. "The adaptation of produced voice-onset time." *Journal of Phonetics* 15, pp. 15-27.

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

Keating, A. 1984. "Phonetic and phonological representation of stop consonant voicing." *Language* Vol. 60 (2), pp. 286-318.

Kewley-Port, D. and M.S.Preston.1974. "Early apical stop production : a voice onset time analysis." *Journal of Phonetics* 2, pp. 195-210.

Klatt, D.H. 1975. "Voice-onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18, pp.686-706.

Klee, T.M., G. Weismer and D.R. Ingrisano. 1976. "Laryngeal timing constraints in plosive-vowel and fricative-vowel syllables, or : is VOT really the best measure of glottal-supraglottal timing?" Paper presented at the 92nd meeting of the Acoustical Society of America, San Diego, CA.

Konefal, J.A. and J. Fokes 1981. "Voice onset time : the development of Spanish/English distinction in normal and language disordered children." *Journal of Phonetics* 9, pp.437-444.

Lisker, L. and A.S. Abramson. 1964. "A cross-language study of voicing in initial stops : acoustical measurements." *Word* 20, pp.384-422.

_________. 1967. "Voice onset time in English stops." *Language and Speech* 10, pp.1-28.

Mack, M. 1983. "Psycholinguistic consequences of early bilingualism : a comparative study of performance of English monolinguals and French-English bilinguals in phonetic, syntactic, and semantic experiments." Unpublished Ph.D. thesis. Brown Univ. Providence, R.I.

MacKay, Ian R.A. 1987. *Phonetics and the Science of Speech Production.* 2nd ed. Boston : Little, Brown and Co.

Major, R.C. 1987. "English voiceless stop production by speakers of Brazilian Portuguese." *Journal of Phonetics* 15, pp.197-202.

Murry, T. and Brown, Jr. W.S.1976. "Peak intraoral air pressures in whispered stop consonants." *Journal of Phonetics* 4, pp.183-187.

Nathan, G.S. 1987. "On second-language acquisition of voiced stops." *Journal of Phonetics* 15, pp.313-322.

Port, R.F., S.Al-Ani and S. Maeda. 1980. "Temporal compensation and universal phonetics." *Phonetica* 37, pp.235-252.

Weismer, G. 1979. "Sensitivity of voice-onset time (VOT) measure to certain segmental features in speech production." *Journal of Phonetics* 7, pp. 197-204.

_________. 1980. "Control of the voicing distinction for intervocalic stops and fricatives :

some data and theoretical considerations." *Journal of Phonetics* 8, pp.427-438.

Williams, L. 1977. "The voicing contrast in Spanish." *Journal of Phonetics* 5, pp.169-84.

_________. 1979. "The perception of stop consonant voicing by Spanish-English bilinguals." *Perception Psychophysics* 21, pp.289-97.

_________. 1980. "Phonetic variation as a function of second-language learning." *Child Phonology and Perception Vol. 2*, pp.185-216.

Winitz, H., C. LaRiviere and E. Herriman. 1975. "Variations in VOT for English initial stops." *Journal of Phonetics* 3, pp.41-52.

Zlatin, M.A. 1974. "Voicing contrast and voice onset time." *Journal of Acoustical Society of America* 56 (3), pp. 981-87.