Kim, M-R, & Lotto, A.J. (2002). An investigation of acoustic characteristics of Korean stops produced by non-heritage learners. In J.J. Ree (Ed.), The Korean Language in America, Vol. 7 (pp. 177-188). American Association of Teachers of Korean.

An Investigation of Acoustic Characteristics of Korean Stops Produced by Non-heritage Learners

Mi-Ran Cho Kim
University of Georgia

Andrew J. Lotto
Washington State University

1. Introduction

Stops are the most common consonants in the systems of the world's languages. In fact, stops are the only consonants which occur in all human languages (Ladefoged & Maddieson, 1996). They also show a great deal of variation in terms of the airstream mechanism, state of the glottis, manner of articulation, and place of articulation. Phonologically, stop consonants are typically classified by the features of [voicing] and [aspiration]. English, for example, has two types of stop that contrast in voicing, while Korean has three types of stop which contrast in degree of aspiration. Previously the nine Korean stops, which are all voiceless word-initially, have been described in several different ways such as aspirated, unaspirated, and glottalized; heavily aspirated, slightly aspirated, and unaspirated; aspirated, lax, and tense, respectively. Since this study focuses on the degree of aspiration in Korean stops, we are going to use heavily aspirated, slightly aspirated, and unaspirated. In both Korean and English, stops occur in three different places of articulation, i.e., bilabial, alveolar, and velar.

The notion of 'interference' from the native language on learning the target language is not new to researchers in the field of teaching second or foreign language. Language learners tend to perceive the target language utterances in terms of the linguistic system of their native language. The difficulties facing a language learner can be predicted by comparing the phonetic realizations of sounds in the native and to-be-learned languages (Best, 1995; Flege, 1995; Holt, Lotto, & Kluender, 1998). This comparison must take place at the sub-phonemic acoustic level. Phonetic research shows that the phonological features of stops do not bear a one-to-one relationship to their physical phonetic realizations. This is partly because distinctive feature theory at the phonemic level uses a discrete binary classification, whereas at the acoustic phonetic level, the feature correlates take on a continuum of values (Zue, 1976). Because of this, in order to soften or eradicate the foreign accent, the language learners must master the phonetic system of the target language (Carr, 1994). Moreover, the comparison of language learner's production with native speakers' production is crucial for identifying areas of difficulty that the language learner may encounter. The minimization of these acoustic differences can serve as a useful goal for teachers.

Acoustic phonetics is a subfield in linguistics in which the actual speech sounds are represented in terms of scientific physical measurements, such as duration, intensity, and amplitude. Although there has been a great deal of research on second and foreign language teaching, rarely is emphasis placed on the importance of precise scientific description of speech sounds produced by the language learners. The present study provides an acoustic description of language learners' productions instead of the typical rather impressionistic description. This study has two purposes: (1) to identify the acoustic characteristics of the word initial Korean stops produced by non-heritage learners of Korean, and (2) to utilize the findings for the improvement of teaching production of Korean stops to non-heritage learners. For the first purpose, three major acoustic variables, closure duration, voice onset time, and fundamental frequency of the following vowel, were measured and the values were statistically tested using ANOVAs to test whether the differences were significant for the distinctions among three types of Korean stops. For the second purpose, the results of the first experiment were compared to the results of previous research where the acoustic characteristics of Korean stops produced by native speakers of Korean were measured. In what follows, an overview of previous research on the acoustic analysis of stop consonants in English and in Korean is presented.

2. Acoustic Characteristics of Stop Consonants

Stop consonants are produced by complete momentary blockage of the vocal tract, followed by release of the airstream. Spectrographically, voiced stop consonants are characterized by the presence of a voice bar during the interval of articulatory closure. The voice bar, which appears as vertical striations, is a band of low frequency periodic energy, which is produced by phonation. The closure interval of voiceless stops is virtually blank like a pause between utterances. Aspirated stop consonants are characterized by long lag between the closure release and the voicing onset of the following vowel (Kent & Read, 1992). The articulatory distinction between types of stops can result in a variety of changes in the acoustics of the sound produced. As these acoustic features are lawfully related to the production differences, they can serve as perceptual cues to the stop distinction. Languages differ in the acoustic cues that are distinctive for stop types. A mature speaker of a language uses all of the language-specific cues to maximize intelligibility.

The most reliable acoustic cue for the distinction of prevocalic English stops is not voicing (during the closure) but rather Voice Onset Time (VOT): the interval between the articulatory release of the stop and the onset of vocal fold vibration of the following segment (Fry, 1979; Kent & Read, 1992). For the so-called voiced stops in English, VOT ranges up to 20 ms, while for voiceless stops, it ranged from 25 ms up to 100 ms (Kent & Read, 1992). Lisker and Abramson (1964) have suggested that the articulatory features of voicing, aspiration, and force of articulation have predictable consequences on VOT. As a result, VOT is sufficient to separate the stop categories of a number of languages. Kim (1994) found that VOT was significantly different between the three types of Korean stops for productions from native speakers. VOT ranged from 71.2 ms to 100.9 ms for heavily aspirated stops, from 14.6 ms to 94.8 ms for slightly aspirated stops, and from 7.0 ms to 22.4 ms for unaspirated stops. Thus, VOT is clearly related to Korean stop type but there is substantial overlap in values across the three stop categories.

Silent closure duration is also acknowledged to be an important acoustic cue to the perception of stop consonant type. Silent closure duration is also known as a major cue for place of articulation: labial stops have a longer closure interval than alveolar or velar stops in natural speech (Repp, 1984). In the productions of native speakers of Korean, closure duration is longest for the unaspirated, shorter for the heavily aspirated, and shortest for the slightly aspirated stops (Kim, 1994). The closure duration ranges from 68.8 ms to 143.2 ms for the unaspirated stops, from 69.8 ms to 140.9 ms for the heavily aspirated stops, and from 43.0 ms to 107.9 ms for the slightly aspirated stops.

The fundamental frequency (f0) of the vowel is also said to vary depending on a voicing distinction of the preceding stop consonants. Keating (1984) suggested that in many languages, the f0 of the vowels following voiceless stops is high, while that following voiced stops is low. For Korean stops, the f0 is highest after the heavily aspirated stops, and higher after the unaspirated stops than the slightly aspirated stops (Kim, 1994).

3. Experiment

3.1 Subjects

Twelve non-heritage learners of Korean participated in this experiment. They were all undergraduate students and enrolled in the intermediate Korean classes at the University of Georgia. Eleven subjects ranged in age between 18 and 21. One subject was 35 years old. All participants had only two semesters of learning experience with Korean. Eleven subjects were native speakers of English and one subject was a native speaker of Japanese.

3.2 Stimuli

For the data collection, a reading list containing six tokens for each of nine stops was prepared. The tokens were all nouns that began with a CV(C) syllable. The vowel was fixed as /a/. In order to measure the closure duration of each stop, the nouns were embedded in a carrier sentence, "i (this) + kes (thing) + i (nominative particle) <u>noun</u> iyeyo (is).", which means "This is <u>noun</u>." The nouns used for the experiment are listed in Table 1.

Table 1. List of Nouns

Stop	Noun
/p^h/	/p^hato/ 'tidal wave', /p^hari/ 'fly', /p^harangse/ 'bluebird', /p^hamun/ 'scandal', /p^hach^{h}o/ 'plantain', /p^hal/ 'arm'
/t^h/	/t^hacho/ 'ostrich', /t^hankwang/ 'coal mine', /t^hal/ 'mask', /t^haakki/ 'percussion instrument', /t^hackaki/ 'typewriter', /t^hawonhyung/ 'oval'
/k^h/	/k^hal/ 'knife', /k^hamera/ 'camera', /k^hangkaru/ 'kangaroo', /k^hanna/ 'canna', /k^halkuksu/ 'noodle', /k^hadû/ 'card'
/p/	/pata/ 'sea', /paü/ 'rock', /pakuni/ 'basket', /pakachi/ 'gourd', /pachi/ 'trousers', /pal/ 'foot'
/t/	/tari/ 'bridge', /tapang/ 'cafe', /tarimi/ 'iron', /taramchü/ 'squirrel', /talre/ 'allium', /tal/ 'moon'
/k/	/kaü/ 'scissors', /kapang/ 'bag', /kapal/ 'wig', /kachi/ 'egg plant', /kamyen/ 'mask', /kayo/ 'folk song'
/pp/	/ppang/ 'bread', /ppalre/ 'washing', /ppalkansek/ 'red', /ppalte/ 'straw', /ppalkengi/ 'communist', /ppalch^{h}isan/ 'Palchi Mountain'
/tt/	/ttangk^{h}ong/ 'peanut', /ttang/ 'earth', /ttalki/ 'strawberry', /ttal/ 'daughter', /ttaoki/ 'crested ibis', /ttam/ 'sweat'
/kk/	/kkach^{h}i/ 'magpie', /kkamakü/ 'crow', /kkangt^{h}ong/ 'can', /kkaktuki/ 'hot pickles', /kkalteki/ 'funnel', /kkatalk/ 'reason'

3.3 Procedure

The recordings were made in a quiet room at the University of Georgia. Each subject was asked to read the list of sentences as naturally as possible at normal speed. In order to facilitate the reading, the list was given to the subjects prior to the recording session. The Korean instructor helped the subjects with practice of the Korean nouns. The experimenter demonstrated the desired speech rate prior to recording. The subjects' speech was recorded on a Sony Digital Audio Tape-corder at a 48-kHz sampling rate. The distance of the microphone (Shure SM48-LC) from the mouth was approximately 4 inches.

3.4 Acoustic Analysis

Six hundred forty eight tokes were analyzed (3 types of stops x 3 places of articulation x 6 tokens x 12 subjects) in terms of (1) closure duration, (2) voice onset time, and (3) fundamental frequency of the following vowel. Measurements were made using the TF32 software (Milenkovic & Read, 2000).

For the closure duration, the period of silence between the last glottal pulse of the nominative particle /i/ and the release of the stop burst was measured. VOT was measured from the stop release to the point at which the waveform shows the first sign of periodicity. For the fundamental frequency of the following vowel, the average fundamental frequency of the first five glottal periods was measured from the sound waveform.

3.5 Statistical Analyses

In order to investigate the characteristics of each of three acoustic variables under three types of stops, a two-way mixed model ANOVA (analysis of variance) was employed.

There were two independent factors in the analyses: one between-subject factor (gender) and one within-subject factor (types of stops). The average values of the six measurements of each acoustic variable were used as the dependent variable for the analysis. After obtaining a significant value of F for the types of stop factor, Holm's sequential Bonferroni procedure was performed at $p<.05$ for the pairwise comparison of the three types of stops (Green, et al. 2000).

3.6 Results

3.6.1 Closure Duration (CD)

Tables 2, 3 and 4 present the mean closure duration for the three types of Korean stops by place of articulation. The closure duration showed a great variation between subjects. The patterns are also not consistent across the place of articulation. For the bilabial and velar stops, the closure duration is longest for the unaspirated stops, shorter for the heavily aspirated stops, and shortest for the slightly aspirated stops. For the alveolar stops, in contrast, the closure duration is the longest for the slightly aspirated stop, shorter for the heavily aspirated stop, and shortest for the unaspirated stop. The analyses indicated that only the differences in velar stops are statistically significant [$F(2,9) = 9.577$, $p = .006$]. Even though the female subjects, in general, produced shorter closures than the male subjects, the gender difference was not statistically significant.

Table 2. Mean and Standard Deviation (SD) of Closure Duration (in ms) for Bilabial Stops

	p^h		p		pp	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	567.8	285.4	519.1	279.0	590.4	371.0
Female	405.0	146.2	355.1	150.6	395.6	162.0
Total	499.9		450.8		509.3	

Table 3. Mean and Standard Deviation (SD) of Closure Duration (in ms) for Alveolar Stops

	t^h		t		tt	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	614.9	327.4	565.2	238.5	588.4	304.0
Female	477.6	281.0	561.7	221.1	431.9	274.8
Total	557.7		563.8		523.2	

Table 4. Mean and Standard Deviation (SD) of Closure Duration (in ms) for Velar Stops

	k^h		k		kk	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	579.0	355.9	503.2	281.5	752.7	376.9
Female	413.0	150.4	422.3	296.1	611.6	332.4
Total	509.8		469.5		693.9	

3.6.2 Voice Onset Time (VOT)

Tables 5, 6, and 7 present the mean VOT for each type of stops by place of articulation. The findings are consistent across three places of articulation. As expected, VOT for the heavily aspirated stops are longer than those of the slightly aspirated stops and much longer than those of the unaspirated stops. VOT ranged from 44.4 ms to 92.0 ms. for heavily aspirated stops, from 31.4 ms to 73.8 ms for slightly aspirated stops, and from 12.4 ms to 51.5 ms for unaspirated stops. The pairwise comparison results indicated that all three differences are statistically significant ($p < .05$). This finding is consistent with previous studies (Lisker & Abramson, 1964; Kim, 1994). Again, even if the mean values of VOT produced by female subjects are consistently longer than those of male subjects, the difference in VOT by gender is not statistically significant.

Table 5. Mean and Standard Deviation (SD) of Voice Onset Time (in ms) for Bilabial Stops

	p^h		p		pp	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	44.4	16.9	31.4	17.8	12.4	2.3
Female	57.6	23.6	62.6	18.3	28.4	11.0
Total	49.9		44.4		19.1	

Table 6. Mean and Standard Deviation (SD) of Voice Onset Time (in ms) for Alveolar Stops

	t^h		t		tt	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	62.1	17.6	39.9	21.6	20.9	11.2
Female	80.2	24.5	44.3	17.0	21.5	5.6
Total	69.7		41.7		21.1	

Table 7. Mean and Standard Deviation (SD) of Voice Onset Time (in ms) for Velar Stops

	k^h		k		kk	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	65.7	15.8	59.0	33.6	43.0	26.1
Female	92.0	13.2	73.8	30.6	51.5	18.8
Total	76.7		65.2		46.6	

3.6.3 Fundamental Frequency of the Following Vowel (f0)

The mean values of fundamental frequency of the following vowel are given in Tables 8, 9, and 10. The findings are not consistent across place of articulation or by gender. For the bilabial stops, the mean value of f0 is highest after the heavily aspirated stops, and higher after the slightly aspirated stops than the unaspirated stops. The differences in f0 for the bilabial stops were not statistically significant. The alveolar stops showed a similar pattern of f0 as the bilabial stops. The f0 was higher after the heavily aspirated stops than after either the slightly aspirated stops or the unaspirated stops. In turn, the f0 after the slightly aspirated stops is higher than that after the unaspirated stops. The pairwise analyses indicated that only the f0 after the heavily aspirated stops differed significantly from the other stop types. For the velar stops, the f0 after the unaspirated stops is slightly higher than after the heavily aspirated stops. In turn, the f0 after the heavily aspirated stops is higher than after the slightly aspirated stops. The pairwise analyses indicated that the f0 differences between the heavily aspirated and the unaspirated are not statistically significant, while the f0 after the slightly aspirated stops is significantly lower than the f0 after the heavily aspirated and the unaspirated.

Table 8. Mean and Standard Deviation (SD) of Fundamental Frequency of the Following Vowel (in Hz)

	p^h		p		pp	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	112.5	11.6	110.3	9.4	107.1	9.4
Female	221.6	9.2	223.3	10.2	226.5	15.7
Total	157.9		157.3		156.9	

Table 9. Mean and Standard Deviation (SD) of Fundamental Frequency of the Following Vowel (in Hz)

	t^h		t		tt	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	114.6	10.3	109.3	8.4	109.7	9.8
Female	227.2	13.1	222.1	15.2	216.0	18.4
Total	161.5		156.3		154.0	

Table 10. Mean and Standard Deviation (SD) of Fundamental Frequency of the Following Vowel (in Hz)

	k^h		k		kk	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	114.4	12.3	110.9	10.6	114.2	12.7
Female	216.7	16.6	212.6	19.6	218.8	23.0
Total	157.0		153.3		157.8	

4. Discussion

The measurements obtained for the non-heritage speakers in this study can be compared to results from native Korean speakers from an earlier study that used the same empirical protocol (Kim, 1994). In that study, all three acoustic features were found to be related to the aspiration distinction for native Korean speakers. The patterns obtained for non-heritage speakers were quite different across the three features.

The mean values of voice onset time, which is considered the most salient acoustic feature for the identification of the stops, differed significantly across the three types of Korean stops for the non-heritage speakers. This result demonstrates that the language learners were using the VOT feature to distinguish the stops. However, the VOTs of the heavily aspirated stops were considerably shorter than measured from the productions of the native Korean speakers (Kim, 1994). The VOTs of both the heavily aspirated and the slightly aspirated stops were similar to the VOTs of the slightly aspirated stops produced by the native Korean speakers. That is, the language learners were using the VOT cue distinctively but in a manner that was clearly not native-like.

The mean closure duration of the present study (355.1 ms to 752.7 ms) were considerably longer than those of native Korean speakers (43.0 ms to 143.2 ms). This was partly due to the slower speech rate, pause, and hesitation typical of new language learners. Closure duration was not related to the distinction between the three types of Korean stops. This is in contrast to the findings of Kim (1994) that native Korean speakers produced a significantly shorter closure duration for slightly aspirated stops.

The clearest example of differences between native and non-native productions was in the relation of f0 to the stop distinction. Non-heritage speakers produced small f0 changes that were not consistently related to stop type. In contrast, Kim (1994) found that native Korean speakers made large f0 changes related to stop type. In fact, the data from native speakers suggest that f0 is the major cue distinguishing slightly aspirated stops from the other two stop types. The average difference in f0 between slightly aspirated and heavily aspirated stops was 53.57 Hz (compared to a paltry 3.17 Hz for the productions of the non-heritage speakers). Clearly, the language learners were not manipulating f0 appropriately for these productions. One reason for this difficulty may be because the f0 pattern in Korean is quite different from the pattern in English productions of stops (low f0 with voiced stimuli and high f0 for voiceless stimuli).

5. Conclusion

The description of non-heritage learner's Korean stop production in terms of acoustic measurements can serve as an excellent guide to Korean instructors. In particular, the results of the current study demonstrate that language learners (predominantly English-speaking) make use of VOT for stop distinctions but fail to use closure duration or f0 in a distinctive manner. These problem areas can be a focus for instructors. Visual representations of the speech sounds, such as spectrograms, are potentially very useful for the language learners, because the language learners often cannot perceive the subtle difference between their productions and those of native speakers. This technique has been used successfully in training Japanese-native speakers to produce the American English /l/-/r/ distinction (Akahane-Yamada, McDermott, Adachi & Takada, 1999). It is hoped that as a first step, an accurate description of the production of Korean learners will help Korean instructors better understand production problems. It is intended to show advantages of acoustic analyses over the previous impressionistic diagnoses of language learner's production, which often can be inaccurate and misleading.

References

Akahane-Yamada, R., McDermott, E., Adachi, T., & Takada, T. (1999). Effect of computer-assisted training on production of English /r/ and /l/ by Japanese. *Journal of the Acoustical Society of America, 106,* 2154.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception of linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167-200). Timonium, MD: York.

Carr, P. (1994). *Phonology.* New York: St. Martin's Press.

Flege, J. E. (1995). Second language learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233-272). Timonium, MD: York.

Fry, D. B. (1979). *The physics of speech*. Cambridge, UK: Cambridge University Press.

Green, S. B., Salkind, N. J., & Akey, T. M. (2000). *Using SPSS for Windows: Analyzing and understanding data* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Holt, L. L., Lotto, A. J., & Kluender, K. R. (1998). Incorporating principles of general learning in theories of language acquisition. In C.M. Gruber, D. Higgins, K.S. Olson, & T. Wysocki (Eds.), *CLS 34: The Panels.* Chicago: *Chicago Linguistic Society, 34*(2), 253-268.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60,* 286-319.

Kent, R., & Read, C. W. (1992). *The acoustic analyses of speech*. San Diego, CA: Singular Publishing.

Kim, M.-R. Cho (1994). *Acoustic characteristics of Korean stops and perception of English stop consonants*. Doctoral dissertation, University of Wisconsin, Madison.

Ladefoged, P. & Maddieson, I. (1996). The sounds of world's languages. Oxford, UK: Blackwell.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20,* 384-422.

Milenkovic, P. H., & Read, C. (2000). CspeechSP [Computer software]. Madison, WI: Authors.

Repp, B. H. (1984). *Closure duration and release burst amplitude cues* (Status Report on Speech Reseach SR-77/78, pp. 137-145). Haskins Laboratories.

Zue, V. (1976). *Acoustic characteristics of stop consonants: A controlled study*. Unpublished doctoral dissertation, MIT, Cambridge.