A perceptual study of the three-way contrast in Korean stops with cross-spliced syllables

Mi-Ryoung Kim
Kmrg@umich.edu

University of Michigan

Abstract

This paper examines the contribution of vocalic information (after the onset of voicing) to the perception of Korean alveolar stops: the aspirated /t^h/, the lenis /t/, and the fortis /t*/. These stops have been analyzed as differing in VOT (Abramson & Lisker, 1964), the glottal width or aspiration (Kim, 1970), and F0 and intensity build-up (Han & Weitzman, 1970). These studies focused on the articulatory and acoustic qualities of the consonants and often assumed that the consonantal portion before the onset of voicing plays the main role in maintaining the three-way distinction. In contrast, the role of the following vowels was given less attention. In order to investigate the contribution of the following vowels, a perceptual study was conducted using stimuli cross-spliced from three naturally produced syllables [t^{h}al] 'mask', [tal] 'moon', and [t*al] 'daughter'. Stimuli were presented to 12 Korean listeners for identification. Each subject responded to a total of 486 tokens. The results show that vowels play the primary role when the cut occurs at the start of voicing. Even with cuts at 10 ms and 40 ms into voicing, the following vowel still plays a clear role. This suggests that vowels carry the important information for distinguishing the three stops.

Introduction

Korean stops are classified into three groups according to the place of articulation, i.e. labials, alveolars, and velars. In initial position, there is a three-way distinction in the manner of articulation. All stops are voiceless in initial position. This is shown in (1):

(1)

	Aspirated	*Lenis*	*Fortis*
labials:	p^{h}ul 'grass'	pul 'fire'	p*ul 'horn'
alveolars:	t^{h}al 'mask'	tal 'moon'	t*al 'daughter
velars:	k^{h}i 'height'	ki 'energy'	k*i 'meal'

Many studies have attempted to clarify the acoustic, physiological, and aerodynamic properties that differentiate the three manner categories (Lisker & Abramson, 1964, 1971; Kim, 1965, 1970; Umeda & Umeda, 1965; Han & Weitzman, 1970; Hardcastle, 1973; Kagaya, 1974; Dart, 1986; Jun, 1990, 1994; Lee, 1990). Lisker & Abramson (1964) acoustically analyzed word-initial stops in several languages, including Korean and found voice onset time (VOT; the temporal relation between stop release and onset of glottal pulsing) to provide a useful measure for differentiating conditions of voicing and aspiration.

Table I presents a summary of previously reported mean VOT values for the three categories of alveolar Korean stops in word-initial position.

	t^h	t	t*	# of speakers
Lisker & Abramson(1964)	96	28	11	2
Kim(1965)	92	38	15	1
Han & Weitzman(1967)	120	28	10	2

Table I. Summary of Mean VOT values for Korean alveolar stops in initial position.

Han & Weitzman (1970) examined the acoustic and perceptual features of the Korean stops. With a listening test using the tape-cutting-and-splicing method, they concluded that both the timing of voice onset and the quality of voice onset (F0 and intensity build-up) are important cues in the phonemic differentiation of Korean stop consonants. That is to say, neither cue by itself seemed able to differentiate the three types of stops.

Abramson & Lisker (1971) investigated the perceptual relevance of the VOT distinction for Korean listeners by presenting a continuum of synthetic VOT variants to native speakers for identification as Korean syllables. Their complicated response patterns and production indicated that although the timing of glottal adjustments relative to supraglottal articulation contributes to the consonantal distinctions, there must be another dimension that works with VOT in distinguishing the categories.

Lee (1990) synthesized the Korean alveolar fortis and lenis stops with the articulatory and acoustic dimensions based on the finding of the previous studies and tested listeners' ability to identify and discriminate the stimuli. The results showed that Korean native speakers did not perceive the stimuli categorically.

Most studies have focused on the differences in the consonant properties. A question that arises looking at these studies is whether the vowel plays a role in distinguishing the three categories. The present research investigates perceptually this issue. To test this, a perception test was conducted using stimuli in which the initial consonant and vowel from naturally produced CVC syllables were cross-spliced. Three outcomes are possible, with different interpretations:

Outcome 1: If consonants carry the main cue, there would be little effect of vowel in the responses to cross-spliced syllables.

Outcome 2: If consonants do not carry the main cue, there would be important effect of vowels in the responses to cross-spliced syllables.

Outcome 3: If neither consonants or nor vowels carry the main cue, there would be a third effect in the responses to cross-spliced syllables.

Method

Stimuli Materials

A 30-year old female native speaker of the Seoul dialect of Korean was recorded on Mac Speech Lab. She read the following utterances: ba t^{h}al ba ba tal ba ba t*al ba where [ba] is a nonsense syllable and /t^{h}al/ 'mask', /tal/ 'moon', and /t*al/ 'daughter' are Korean words (see Fig. 1). Each utterance was read several times.

Stimuli were digitized at 10 kHz and edited using a wide-band spectrogram and waveform displays. From the multiple repetitions, nine stimuli (3 syllables x 3 tokens) were chosen. Figure 1 illustrates three sample stimuli among the nine.

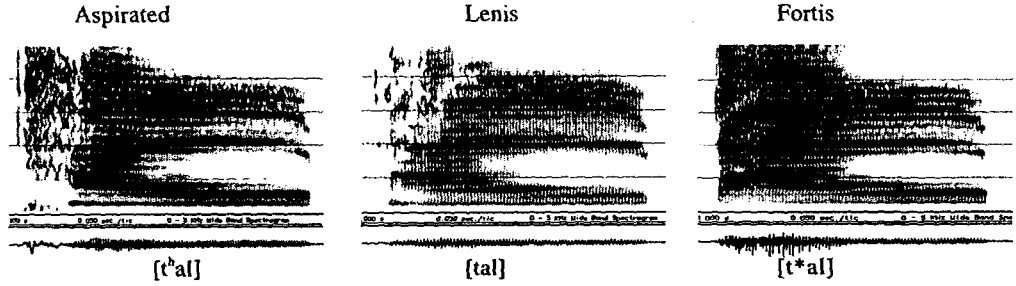

Figure 1. Wide-band spectrograms and waveforms of the three alveolar stops.

The consonantal shapes of the three stimuli are well differentiated in the spectrogram: the aspirated stop has a large amount of aspiration, the lenis stop less aspiration, and the fortis stop no aspiration. The vowels following the three stops also show some differences: the fortis stop has a

greater intensity build-up at the onset of voicing than the aspirated and lenis stops. For the lenis stop, a mixture of voicing and noise at the onset of the vowel is seen, as noted in earlier studies (Kim, 1970; Han & Weitzman, 1970; Lee, 1990). Also, at voice onset, the aspirated and lenis stops, but not the fortis stop, exhibit formant cutback. The other six stimuli employed here have similar spectral characteristics.

Cross-splicing techniques were used to investigate the role of the vowel in perception of the consonantal contrast. The original 9 stimuli were first divided into onset and rime (using spectrographic and waveform displays), defined in 3 ways depending on the location of the cut: (a) at onset of voicing for the vowel (0ms); (b) 10 ms into vocalic voicing; and (c) 40 ms into vocalic voicing. These 3 methods are illustrated in Fig. 2 with a single token of /t^hal/.

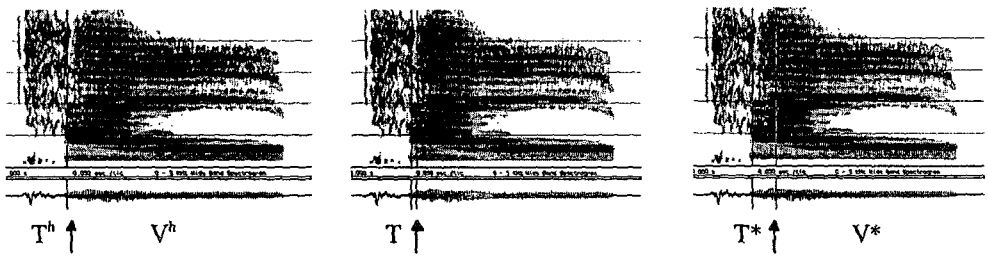

Figure 2. Wide-band spectrograms and waveforms of the three alveolar stops with three different cut locations.

Next, onsets and rimes were cross-spliced across the three consonant types, as shown in Table II (where T^h and V^h =aspirated onset and rime, T and V =lenis onset and rime, and T* and V* =fortis onset and rime).

	V^h	V	V*
T^h	T^hV^h	T^hV	T^hV^*
T	TV^h	TV	TV*
T*	T^*V^h	T*V	T*V*

Table II. Cross-spliced syllables.

Stimuli along the diagonal (i.e., T^hV^h, TV, and T*V*) were also spliced from different tokens of the same consonant type, so that no original stimulus was a test stimulus. For each way of cut, 18 onset and rime pieces (3 sets of 6 each) were obtained which were then cross-pasted to make 27 new syllables (3sets of 9 each). That is to say, each original syllable yielded nine test syllables, for a total of 81 test syllables.

Six randomizations of the 81 test stimuli were created for a total of 486 test items. The intersimulus interval was 2 s and the interblock interval (between blocks of 81 stimuli) was 10s.

Subjects and Procedure

All test sequences were output to digital audio tape and were presented over headphones to listeners. Listeners were asked to identify the syllable which they thought they heard and to guess on answers which were not clear. Answer sheets with the Korean letters for the 3 different syllables were provided; listeners responding by circling the appropriate letter. Each listening session was over a half hour.

Listeners were 12 adult native speakers of Korean (6 males and 6 females). Their ages ranged from the late twenties to the early thirties, with a mean age of 30.6 years. Subjects showed the following dialect varieties: 9 speakers of the Seoul dialect, 2 of the Kyungsang dialect, and 1 of the Chonnam dialect. Both speakers of the Kyungsang dialect seemed to have trouble in producing

the three stops, but could distinguish the three syllables from the tape. All subjects were living in the U.S. at the time of listening test. Four subjects had lived in the U.S. less than a year, four less than three years, and the remaining four more than 5 years, with a mean of 3.1 years. None of the subjects had any formal linguistic training and none reported having hearing or speaking deficiencies.

Results

Average responses of 12 listeners to the cross-spliced syllables at 0 ms are given in Figure 3.

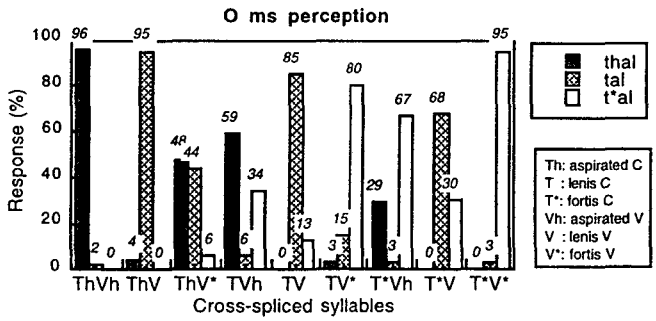

Figure 3. Average responses of 12 native speakers of Korean to the cross-spliced syllables at 0 m.

The vertical axis shows the percentage of the responses to each cross-spliced syllable and the horizontal axis nine cross-spliced syllables. The three different bars represent the three choices. Of the six syllables that contained conflicting cues, four, T^hV, TV^h, TV*, and T*V, were perceived according to the vowel and two, T^hV^* and T^*V^h, according to the consonant. Thus, the cases where the vowel had greater influence out-numbers the cases where the consonant had greater influence.

At 0 ms cut, the three-way contrast was realized by the differences of vowels rather than those of consonants. For TV^h, TV, and TV*, where the vowels are differentiated, the responses of these cross-spliced syllabes showed the three contrasts. However, the response of T^hV, TV, and T*V, where the consonantal portions were differentiated, showed the two contrasts only. This indicated that the three-way manner distinction was largely led by the differences of vowels.

Figure 4 shows that responses to the cross-spliced syllables at 10 ms depend on the vowel.

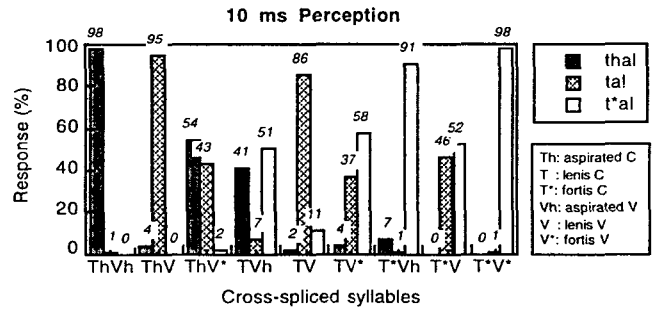

Figure 4. Average responses of 12 native speakers of Korean to the cross-spliced syllables at 10 ms.

Han and Weitzman (1970) suggested that both the timing of voice onset and the quality of voice onset are important cues in the phonemic differentiation of Korean stops. In the data where each consonant contained 10 ms voicing, one would expect that the responses to the cross-spliced syllables depend on the consonant. However, as shown in Figure 4, T^hV, TV^h, and TV* were mainly influenced by the vowel, although T^hV^* and T^*V^h by the consonant. In the case of T^hV^*, 43 percent of the responses were influenced by neither consonant nor vowel. For T*V, half of the responses were influenced by the vowel and the remaining half by the consonant. Therefore, the approximate ratio between the influences of T and V is 2.5 to 3.5.

Figure 5 illustrates the average responses to the cross-spliced syllables at 40 ms.

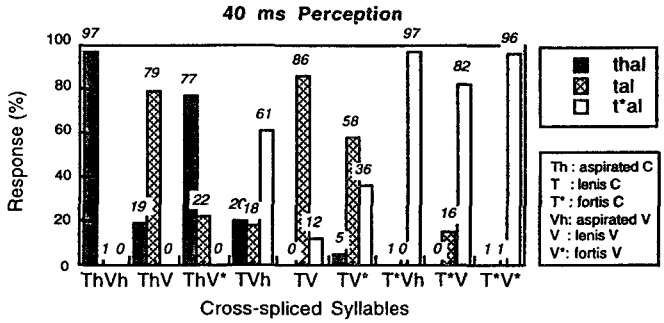

Figure 5. Average responses of 12 native speakers of Korean to the cross-spliced syllables at 40 ms.

Even when the consonant portion contained 40 ms voicing, the response to T^hV still mainly depended on the vowel, although the response to the others on the consonant. In the case of TV*, there was an equal amount of T and V* influence. The ratio between the influence between T and V was approximately 4.5 to 1.5.

In summary, for stimuli cut at 0 ms, perception was guided by the vowel. For stimuli cut at 10 ms into voicing, the influence of the vowel and the consonant was almost equal. And for the stimuli cut at 40 ms into voicing, the influence of the consonant was greater. However, in this case, the vowel still played a role. In general, the cut timing into vowels increases, the role of vowels in the three-way contrast becomes weaker. That is to say, the less the vowel was left on the consonants, the less accurately the consonant was perceived.

Discussion and Conclusion

This paper examined the role of vowels in the perception of the three-way contrast in Korean stops and conducted the perception test using the cross-spliced syllables to see whether the vowels after the onset of voicing carry a perceptual cue in maintaining the contrasts. The results suggest that vowels play a clear role in maintaining of the three-way distinction in Korean stops.

The stimuli used in the present study showed clear acoustic differences in VOT and F0 at the start of voicing as noted by previous studies (Lisker & Abramson, 1964; Han & Weitzman, 1970). However, such cues are clearly not sufficient for the perception of the three-way contrast in Korean stops.

While most subjects showed similar responses, some subjects showed peculiar responses. For example, for the response to T^hV^*, no one heard it as the fortis except a female speaker of the Chonnam dialect. However, she responded it as the fortis about 70 percent. Jun (1990) reported that, in the Chonnam dialect, the fortis or aspirated stop precedes a vowel, it is realized with a low or rising tone depending on its length. As noted by Jun (1990), a speaker of the Chonnam dialect mainly depended on the pitch of the vowel. For the response to T^hV, however, most subjects

seemed relying on the pitch of vowel. This implies that, regardless of the dialectal differences, the perception of speakers could be influenced by the vowel.

In the present study, a systematic account of the relationship between our percetual result and other specific acoustic parameters such as the formant transitions, total duration of the syllable, vowel duration, vowel mean amplitude, and vowel peak amplitude has not been taken. For a future stduy, we could investigate what kind of vowel information affords the listeners' perception of the Korean contrasts.

In summary, using the cross-spliced techniques, this study investigated whether the vowels after the onset of voicing play a role in maintaining the contrasts. The results of the perception test could be summarized as follows: vowels play the primary role in distinguishing the three stops for cross-spliced syllables at 0 ms cut. Even the cuts made at 10 ms and 40 ms, vowels still played a clear role in the perception of the three-way distinction. Overall, the perceptual data in this study suggest that vowels must be considered as carrying important perceptual cues in distinguishing the three Korean categories.

Acknowledgment

I would like to thank Prof. Pam Beddor and San Duanmu for their valuable comments and 12 subjects for serving as subjects in the experiments. I am also grateful to James Harnsberger for assisting in the Lab procedures.

*The original version of this paper was presented at Formal Linguistic Society of Mid-America VII (FLSMVII), The Ohio State University, Columbus, 17-19 May 1996.

REFERENCES

Abramson, A. S. and L. Lisker. 1971. "Voice Timing in Korean Stops", *Haskins Laboratories Status Report on Speech Research* 27: 179-184.

Abramson, A. S. and L. Lisker. 1972. "Voice-timing perception in Spanish word-initial stops", *Journal of Phonetics*. 1: 1-8.

Borden, G. J. and K. S. Harris. 1984. *Speech Science Primer*. Baltimore: Williams & Willkins.

Dart, S. N. 1987. "An Aerodynamic Study of Korean Stop Consonants: Measurements and Modeling", . *J. Aoust. Soc. Am.* 81(1): 138-147.

Han, M. S. and R. S. Weitzman. 1967. " Acoustic Features in the Manner Differentiation of Korean Stop Consonants" *Studies in the Phonology of Asian Languages* (USC, Acoustic Phonetics Research. Lab. ,Los Angeles), 5.

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

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

Hirose, H., Lee, C. Y. and T. Ushijima. 1974. "Laryngeal Control in Korean Stop Production", *Jouranl of Phonetics* 2: 145-152.

Jun, S. A. 1990. "The Accentual Pattern and Prosody of the Chonnam Dialect of Korean", *OSU Working Papers in Linguistics* 38: 121-140.

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

Kim, C. W. 1965. "On the Autonomy of the Tensity Feature in Stop Classification (With Special Reference to Korean Stops)", *Word* 21(3): 339-59.

Kim, C. W. 1970. "A Theory of Aspiration", *Phonetica* 21: 107-116.

Lee, S. H. 1990. "Korean Lenis and Fortis stops: Synthesis and Categorical Speech Perception Task", *OSU Working Papers in Linguistics* 38: 105-120.

Lisker, L. and A. S. Abramson. 1964. "A Cross-Language Study of Voicing in Initial Stops: Acoustical measurements", *Word* 20(3): 384-422.