VOICE TIMING IN KOREAN STOPS*

ARTHUR S. ABRAMSON AND LEIGH LISKER

Linguists have disagreed over the features distinguishing the three manner categories of Korean plosives. The three categories of labial, apical and dorsal stops and palatal affricates are variously described for initial position, using one or more of the following terms:

I. voiceless, tense, long, and glottalized

II. voiceless, lax and slightly aspirated

III. voiceless, heavily aspirated, and lax by some but tense by others

A further complication is the frequent voicing of Category II in a medial voiced environment.

We have devoted much of our research effort to questions of laryngeal control in stop consonants. We have shown that various conditions of voicing and aspiration in word-initial stops in a wide variety of languages depend upon differences in voice onset time (VOT), the temporal relation between stop release and onset of glottal pulsing (Lisker and Abramson 1964). Some aspects of the conflicting descriptions of Korean plosives suggested that we test the efficacy of VOT in that language. Combining data from our 1964 study with some recent additions, we present VOT measurements for two native Korean speakers' initial apical stops in Figure 1. The abscissa shows VOT in intervals of 10 msec.; zero represents the moment of stop release. The ordinate shows the frequency distribution of VOT values for each of the three categories. Although Speaker B tends to have slightly higher values, the overall results are quite comparable for the two speakers and for the labial and dorsal stops not shown in the figure. Category III is well separated from the others, but I and II overlap somewhat. Similar data have been published by others (Han and Weitzman 1970, Kim 1965). Of course, where II assimilates to preceding voicing in medial position, VOT separates all three categories.

The foregoing mixed results made us wonder to what extent VOT might provide sufficient perceptual cues for discriminating the three categories. Also, having shown

* This research was supported mainly by the National Institute of Child Health and Human Development and the National Institute of Dental Research.

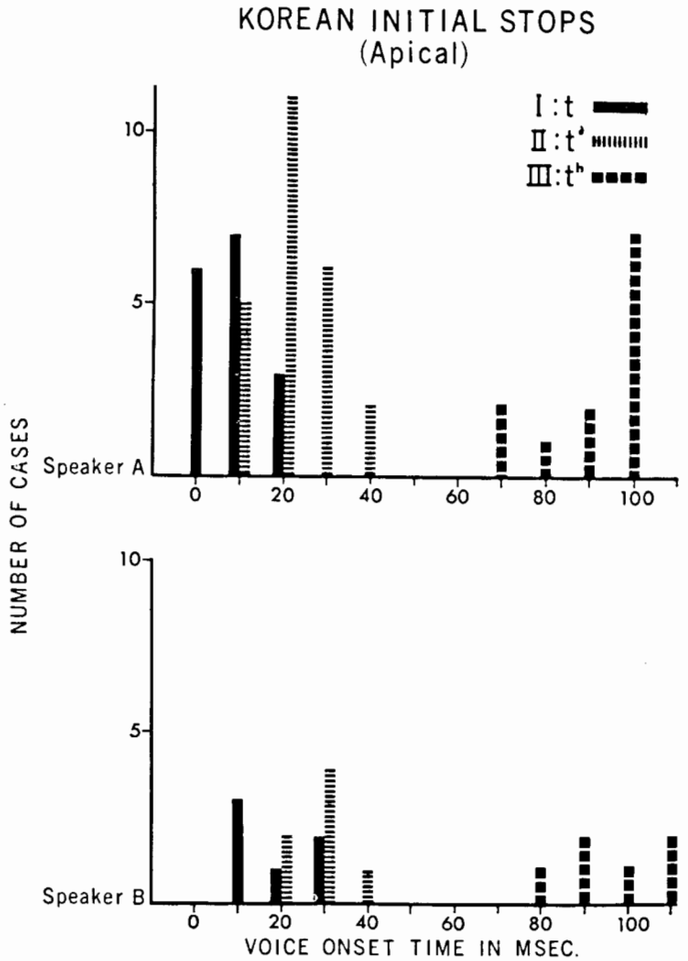

Fig. 1.

the perceptual efficacy of VOT for Thai, Spanish and English (Lisker and Abramson 1970) we wished to extend our comparative phonetic investigation of the dimension to Korean by studying perception as well as production. Lest we later find instability in the phonological distinctions of concern to us, we proved that randomized words differing only in inititial stop categories could be identified with ease. We then exposed native speakers to a continuum of synthetic VOT variants ranging from a voicing lead of 150 msec. before the release of the stop to a voicing lag of 150 msec. after the release, for identification as Korean syllables at each

place of articulation. There were two experimental conditions: (1) a restricted range with all voicing lead variants excluded, thus apparently simulating spoken Korean; (2) the full continuum, thus including variants found only in non-initial position in the spoken language. The range is divided into 10 msec. steps except for the portion from a lead of 10 msec. to a lag of 50 msec., which is divided into 5 msec. steps.

We present labelling responses for the synthetic apical stops only, but the data are typical of all three places of articulation. Figure 2 contains the identification curves for the restricted range. Values of voicing lag are indicated along the abscissa and percent identification as each stop along the ordinate. The left half of the figure is blocked out to show that no lead variants were used. The five subjects responded to the stimuli in three ways. At the top of the figure we see that HL called the variants from 0 to 50 msec. Category I and the rest, Category III; he heard none as II. The middle display shows a partition of the range into I, II and III, in that order; these three subjects, then, behaved much as if VOT were a straightforward cue. At the bottom of the figure, YH divides most of the stimuli between II and III, while weakly favoring I only at 60 and 70 msec.

The responses to the full continuum, including the lead variants marked with negative VOT values, are given in Figure 3. Three response patterns are shown by the four subjects. At the top of the figure, BC simply divides the range into I, II and III, but with occasional labelling of lead variants as II. By and large, she would seem to be hearing voicing lead as a badly pronounced version of the unaspirated stop. We can perhaps understand her vacillation by looking at the middle of Figure 3. There we see two subjects who yielded the startling result that only variants with voicing lead were heard as II, while the rest of the continuum was divided between I and III. It should be recalled that audible laryngeal pulsing does not occur during initial stop occlusions in Korean; therefore, the obvious interpretation of our data is that upon detecting such abnormal voicing at least some Koreans feel they must assign it to the one category that has voiced occlusions in any context at all. This implies that they are somehow aware of glottal pulsing, or the underlying laryngeal gesture, as a component of II. At the bottom of Figure 3, CH not only does the same thing but also assigns several slightly aspirated variants — those from 35 to 80 msec. of lag — to Category II.

The complicated response patterns and production data lead us to two inferences: (1) the timing of glottal adjustments relative to supraglottal articulation contributes to the Korean distinctions, and (2) there must be another dimension that works with VOT in distinguishing the categories. An accumulation of acoustic data on the matter has been furnished by Han and Weitzman (1970) and Kim (1965), in addition to Kim's physiological data (1965, 1970). We are tempted to believe that the difficult question of the distinction between Categories I and II in initial position will be resolved by further examination of laryngeal mechanisms. Recent fiberoptics work by R. Kagaya supports this belief (Kagaya 1971). Also, some speakers have quite

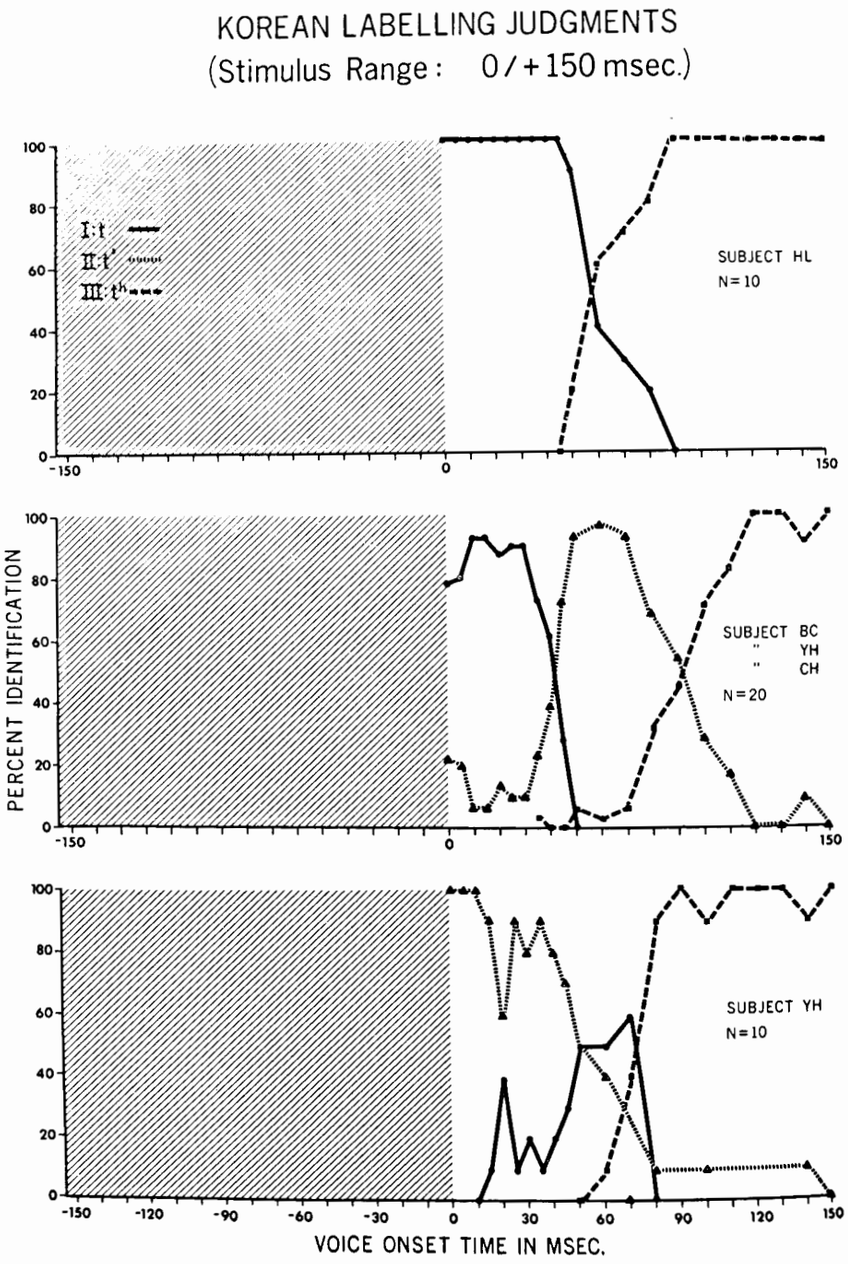
KOREAN LABELLING JUDGMENTS
(Stimulus Range: 0/+150 msec.)
I:t
II:t'
III:t^h
SUBJECT HL
N=10
SUBJECT BC
" YH
" CH
N=20
SUBJECT YH
N=10
PERCENT IDENTIFICATION
VOICE ONSET TIME IN MSEC.

Fig. 2.

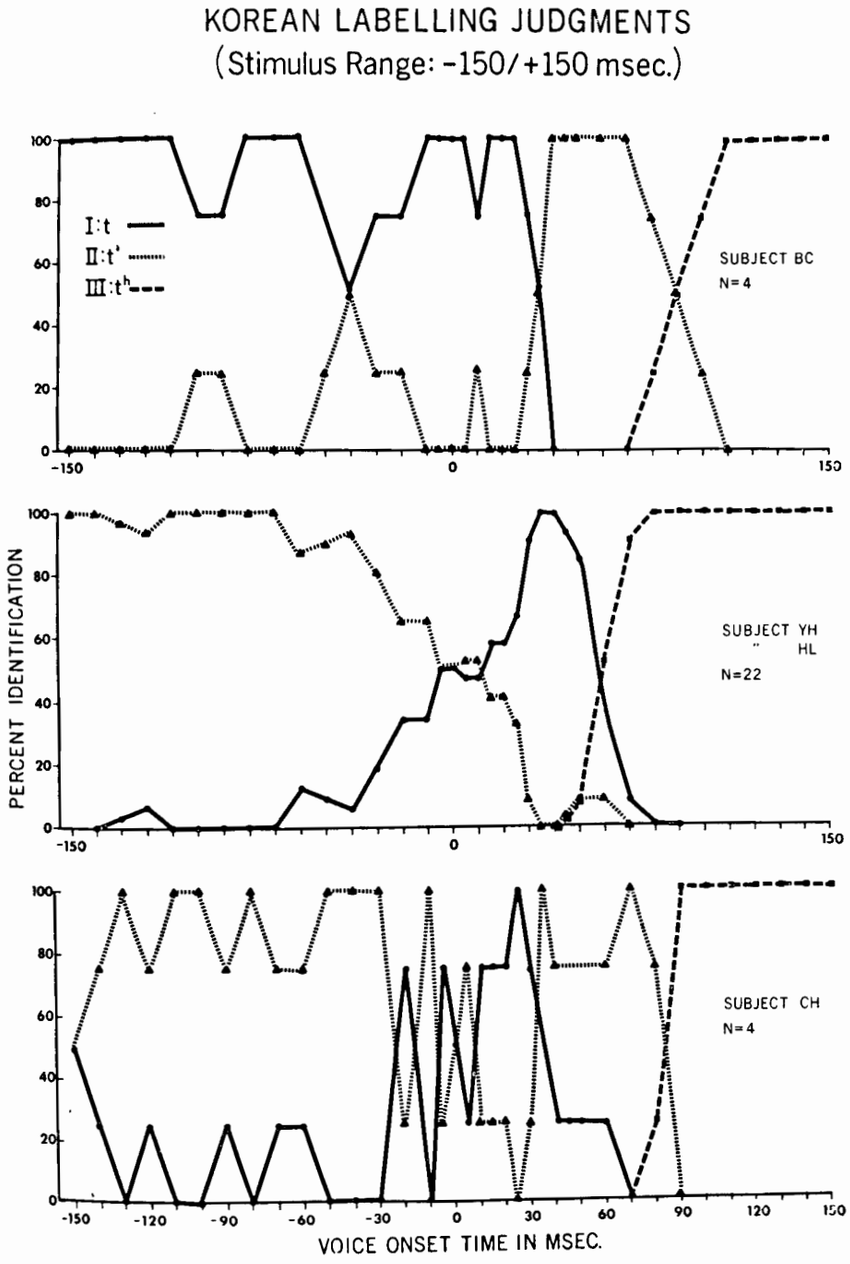
KOREAN LABELLING JUDGMENTS
(Stimulus Range: -150/+150 msec.)
I:t
II:t'
III:t^h
SUBJECT BC
N=4
SUBJECT YH
" HL
N=22
SUBJECT CH
N=4
PERCENT IDENTIFICATION
VOICE ONSET TIME IN MSEC.

Fig. 3.

audible vocal fry or laryngealization in Category I. We plan to take a close look at this phenomenon by means of our fiberoptics system.

Haskins Laboratories, New Haven, Connecticut
also Abramson, University of Connecticut, Storrs
and Lisker, University of Pennsylvania, Philadelphia

REFERENCES

Han, M.S. and R.S. Weitzman
1970 "Acoustic Features of Korean /P,T,K/, /p,t,k/ and /p^h,t^h,k^h/", *Phonetica* 22:112-128.
Kagaya, R.
1971 "Laryngeal Gestures in Korean Stop Consonants", *Research Institute of Logopedics and Phoniatrics, University of Tokyo, Annual Bulletin* 5:15-23.
Kim, C-W.
1965 "On the Autonomy of the Tensity Feature in Stop Classification (With Special Reference to Korean Stops)", *Word* 21:339-359.
1970 "A Theory of Aspiration", *Phonetica* 21:107-116.
Lisker, L. and A.S. Abramson
1964 "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements", *Word* 20:384-422.
1970 "The Voicing Dimension: Some Experiments in Comparative Phonetics", *Proceedings of the VIth International Congress of Phonetic Sciences*, Prague 1967, (Prague, Academia) pp. 563-567.

DISCUSSION

LINDQVIST (Stockholm)
It seems to me that there are many acoustic features that distinguish the type I 'hard' stops from the type II 'weak' stops and the relative importance of these varies from dialect to dialect. These acoustic features may be voicing lag respiration aspiration, quality of the voicing in the following vowel, fundamental frequency and type of voice onset. The dialectal backgrounds of the subjects may then explain the different responses.

Fiberoptic examination of the larynx as well as intra-aural pressure measurements have revealed that the 'hard' stops are produced with a glottal abduction gesture DURING the occlusion, while the 'weak' stops are produced with a glottal abduction gesture which has its maximum at the RELEASE of the stop. The vocal folds are for the latter case in voicing position about half of the occlusion time. At least for the subject studied, it seems as if the lack of voicing for the 'weak' stops in utterance initial position is accidental in the sense that no extra effort is made to make them voiced and the probability for voicing to start is very small, as the vocal folds are in voicing position only in the beginning of the stop. Maybe this can explain why the subjects accepted the voiced initial stops even if these are hardly found in real speech.

ABRAMSON
I trust Mr. Lindqvist did not really mean to say that the absence of voicing during the closure of Korean class II stops in initial position is "accidental". Although the glottal adjustment may not be different from that of the same phonological class in intervocalic position, the regular voicelessness initially cannot be termed 'accidental'.

GRAHAM STUART (Silver Spring, Md.)
Recently, one of my pupils, Dr. Ikmo Lee, has defended a still unpublished dissertation dealing with the perception of vocal source features as consonantal co-articulations. He devoted considerable attention to this problem of the Korean stops and presents experimental data and conclusions very much in agreement with yours and those of other experimenters. In discussions with me however, Lee suggested as a personal insight, rather than as an experimental conclusion, a hypothesis (similar to an earlier proposal of Kim) to explain the incomplete separation of stop categories I and II by voice onset time cues alone. He proposed, namely, that Korean has two types of stop closure, which we may call fortis and lenis. Fortis closure is firm and fully capable of containing the vocal tract pressure; explosion occurs when the articulators are suddenly separated. Lenis closure is weaker and less effective and explosion may occur when the pressure of the contained air reaches the point at which it can blow the articulators apart, just as the vocal chords are forced apart by the subglottal air in phonation, or is, in any case, facilitated by the contained air pressure. The lenis stops are exploded in initial position with only the delay in voice onset time necessary to permit the vocal tract pressure to facilitate or effect the explosion — with the resulting weak aspiration that we know. In medial position, continued phonation is the pressure reducing device used to prevent the critical build-up of vocal tract pressure before the end of the normal occlusive interval. In the apical stops, the fortis occlusion seems to be made with an apical thrust, the lenis occlusion with a coronal lifting of the tongue tip as in the English dental fricatives but against the alveolar ridge instead of the teeth. The difference in pressure exerted by these two different articulators can be sensed by the finger inserted into the mouth.

The Korean stops would, then, be distinguished by the intersection of two bivalued parameters, FORTIS ~ LENIS and ASPIRATE ~ NON-ASPIRATE, rather than one three-valued parameter, NON-ASPIRATE ~ PARTIALLY-ASPIRATE and FULLY-ASPIRATE. Category I would be non-aspirate and fortis, Category II, non-aspirate and lenis, and Category III, aspirate and fortis.

Of course, more and clearer experimental evidence is needed to assert this hypothesis with confidence.

ABRAMSON
Ikmo Lee's insight, as transmitted by Professor Stuart, appears to be a special version of the arguments, supported in part by data, advanced by C-W. Kim (1965), even though Kim himself in a more recent statement on Korean stops (1970) chooses not

to refer to those conclusions. Concerning Lee's speculation as to the inability of a "lenis closure" to contain the buccal air pressure likely to be reached during a Korean stop occlusion, the analogy with the response of the vocal folds to subglottal air pressure is highly improbable. An occlusion of the vocal tract that is sufficiently complete and firm to qualify as a genuine stop closure almost certainly fails by far to fulfill the aerodynamic-myloelastic requirements for a sudden passive release of the articulator. Of course, if the occlusion really were so loose and weak, we might expect to hear frequent productions of affricates and even fricatives instead of the attested stops; that is, I refer here to the phenomenon of lenition. By the way, why should there be a "pressure reducing device used to prevent the critical build-up of vocal tract pressure before the end of the normal occlusive interval" in medial position but not in initial position?