Acoustic and aerodynamic correlates of Korean stops and fricatives

Taehong Cho, Sun-Ah Jun and Peter Ladefoged

*Department of Linguistics, University of California, Los Angeles, 405 Hilgard Ave.,
Los Angeles, CA 90095–1543, U.S.A.*

This study examines acoustic and aerodynamic characteristics of consonants in standard Korean and in Cheju, an endangered Korean language. The focus is on the well-known three-way distinction among voiceless stops (i.e., lenis, fortis, aspirated) and the two-way distinction between the voiceless fricatives /s/ and /s*/. While such a typologically unusual contrast among voiceless stops has long drawn the attention of phoneticians and phonologists, there is no single work in the literature that discusses a body of data representing a relatively large number of speakers. This study reports a variety of acoustic and aerodynamic measures obtained from 12 Korean speakers (four speakers of Seoul Korean and eight speakers of Cheju). Results show that, in addition to findings similar to those reported by others, there are three crucial points worth noting. Firstly, lenis, fortis, and aspirated stops are systematically differentiated from each other by the voice quality of the following vowel. Secondly, these stops are also differentiated by aerodynamic mechanisms. The aspirated and fortis stops are similar in supralaryngeal articulation, but employ a different relation between intraoral pressure and flow. Thirdly, our study suggests that the fricative /s/ is better categorized as "lenis" rather than "aspirated". The paper concludes with a discussion of the implications of Korean data for theories of the voicing contrast and their phonological representations.

© 2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

In most of the world's languages, when stops have a three-way contrast, they are categorized simply in terms of voicing and aspiration: "voiced", "voiceless unaspirated" or "aspirated" (e.g., Lisker & Abramson, 1964; Klatt, 1975; Keating, 1984; Ladefoged & Maddieson, 1996). Korean stops, however, are typologically unusual in that they have a three-way contrast, but they are all *voiceless* in word-initial position. They also use the same airstream mechanism, i.e., pulmonic egressive. The three different categories are often called lenis, fortis, and aspirated, and each of these occur at three places of

Address correspondence to T. Cho. Present address: Max Planck Institute for Psycholinguistics, Postbus 310, 6500 AH, Nijmegen, The Netherlands. E-mail: taehong.cho@mpi.nl

TABLE I. Obstruents in Korean: lenis, fortis, and/or aspirated distinctions

Lenis series	p	t	t∫	k	s
Fortis series	p*	t*	t∫*	k*	s*
Aspirated series	pʰ	tʰ	t∫ʰ	kʰ	

TABLE II. Minimal contrasts for Korean obstruents in word-initial position

Lenis			Fortis			Aspirated	
paŋ	"room"	p*aŋ	"bread"	pʰaŋ	"bang"		
tal	"moon"	t*al	"daughter"	tʰal	"mask"		
t∫ata	"to sleep"	t∫*ata	"to squeeze"	t∫ʰata	"to kick"		
kæta	"to fold up"	k*æta	"to break"	kʰæta	"to dig up"		
sata	"to buy"	s*ata	"to wrap"	—	—		

articulation: bilabial, denti-alveolar, and velar. The lenis stops have been described as lenis, breathy, and slightly aspirated, the fortis stops as tense, laryngealized, and unaspirated, and the aspirated stops as being strongly aspirated.

Along with the stops, Korean has a three-way contrast in affricates and a two-way contrast in denti-alveolar fricatives, which we will call *plain* /s/ *vs.* fortis /s*/, so as to avoid any unnecessary confusion brought about by the terminology *lenis vs. aspirated*. Table I shows all the Korean obstruents. In this table, /s/ is categorized as lenis for the sake of simplicity. The diacritic "*" is used to mark fortis obstruents. Table II gives minimal triplets (or, in the case of the denti-alveolar fricatives, a minimal pair) for the obstruents in Table I.

As reviewed in the next section, several studies have examined a number of acoustic/aerodynamic parameters that might differentiate the three-way contrast in stops. Nevertheless, as far as we know, there is no single work documented in the literature that addresses all these parameters with a relatively large subject pool. Most previous studies have data from one to six speakers. The most comprehensive is Dart's (1987) study, which had data from 10 speakers, but was limited to aerodynamic data for the lenis and fortis stops, excluding the aspirated stop. To understand the phonetic properties underlying the three-way contrastive stops in Korean, we need to collect a body of data from a relatively large number of speakers and examine a variety of phonetic parameters at the same time in much the same way. Without such controls, we cannot be assured about the generalizability of the data. It may be possible that any observed differences can be simply due to the procedures used to obtain data or due to characteristics specific to the small number of speakers recorded. In the present study, we collected both acoustic and aerodynamic data from 12 speakers in two dialects (eight speakers from Cheju Korean and four from Seoul Korean), and examined a variety of phonetic parameters at the same time. Although the size of the sample in the present study may not be large enough to make generalization about the entire population, it is hoped that the present study, which provides ample data from various phonetic parameters, will serve as an improved reference for future research related to Korean obstruents and fine-grained modeling of speech production.

We have some additional specific goals in mind. Firstly, we will examine not only the acoustic properties of each consonant (e.g., VOT, burst energy, and closure/frication

duration) but also properties of the following vowel such as voice quality and fundamental frequency. In particular, we will investigate how and to what degree the three-way contrast of stops can be differentiated by the voice quality of the following vowel.

Secondly, we will examine the aerodynamic properties of stops in terms of airflow and air pressure. There are only a few studies available that have examined airflow and air pressure of Korean stops (e.g., Dart, 1987; Silverman & Jun, 1994). As far as we know, the current study is the first to examine the airflow and air pressure for all three stop categories in Korean.

Thirdly, we will investigate the acoustic properties of the two contrasting fricatives, /s/ and /s*/. Unlike stops, the phonetic properties of the fricatives have not been well studied, and the categorization of the plain /s/ has been controversial. It is sometimes categorized as lenis and sometimes as aspirated (cf. Kagaya, 1974; Iverson, 1983*a*; Jun, 1993; Park, 1999). By examining the acoustic properties of fricatives using the same methods as for stops, we will determine as to what acoustic parameters differentiate /s/ and /s*/, and how the plain /s/ is best classified. The present study is the first one directly comparing the acoustic properties of stops with those of fricatives. Based on this comparison, we will provide phonetic evidence that the plain /s/ belongs to a lenis-like category.

Fourthly, this study provides a cross-dialectal comparison across two very different dialects, Seoul and Cheju Korean. Seoul Korean is the standard dialect of Korean spoken in Seoul, and Cheju Korean is spoken on Cheju Island, located south of the Korean Peninsula and west of Japan. Cheju is the most conservative dialect in Korean in that it still preserves a Middle Korean vowel, the back-mid-unrounded vowel. This vowel was merged with the back-mid-rounded vowel in all other dialects of Korean around the late 1800s. Cheju shares the same morphosyntactic system as other dialects of Korean, and has been traditionally considered as simply a dialect of Korean. On the other hand, due to its having a significant number of different lexical items, especially suffixes and verbal endings, Cheju is not mutually intelligible with the rest of Korean, and might be considered to be a separate language. In this paper, however, we will assume that Cheju is, from a phonetic viewpoint, simply a dialect of Korean since, as mentioned earlier, the Cheju sound system is the same as that of standard Korean except for the back-mid-unrounded vowel. Our own previous work has shown that the Seoul and Cheju dialects maintain the same consonantal phonological systems (Cho, Jun & Ladefoged, 2000). The current paper will allow us to examine as to how the phonetic properties of the consonants are different or similar in these two extreme Korean dialects (Cheju as the most conservative and Seoul as the most dynamic in terms of language change).

Another merit in examining the Cheju dialect lies in documenting an endangered language. Due to the high level of education on Cheju Island and the influence of mass media, it is not easy to find pure native speakers of Cheju in the main city. Presently, it is spoken largely by people over 60, living in the inland, in rural areas. With the help of local Cheju dialect experts, we were able to find several native speakers of Cheju in the hinterlands of the island, all of whom still spoke Cheju, uninfluenced by standard Korean. Thus, we believe that documentation of the phonetic properties of Cheju obstruents in comparison with those of standard Korean has its own merit.

Finally, we will consider as to how the Korean data bear on theories of the voicing contrast and how the phonological representation or featural specification of the three-way contrastive Korean stops can be linked to their phonetic correlates. It is suggested

that the current models of speech production are not specific enough to characterize language-specific phonetic realizations. In order to build finer-grained speech production models, we need to capture phonetic details that are not governed by universal phonetic implementation rules but by language-specific phonetic rules that have to be stated in the grammar of the language.

1.1. *Previous investigations of Korean obstruents*

1.1.1. *Stops*

Researchers have proposed different ways of distinguishing Korean stops in word- or phrase-initial positions. The mean VOT distinguishes the three-way contrast, but the VOT ranges overlap (Lisker & Abramson, 1964; C.-W. Kim, 1965, 1970; Han & Weitzman, 1970; Hardcastle, 1973; Hirose, Lee & Ushijima, 1974; M.-R.C. Kim, 1994; Y. Kim, 1995; J.-I. Han, 1996; Cho, 1996). In general, VOT in phrase-initial position (or in word-initial position in isolation) is shortest for the fortis stop, intermediate for the lenis stop, and longest for the aspirated stop. While these stops are all voiceless phrase-initially (or word-initially in isolation), the lenis stop is voiced intervocalically in word-medial or phrase-medial positions (Cho, 1990; Jun, 1993, 1995), in a phenomenon known as Intervocalic Lenis Stop Voicing. That is, lenis stops assimilate phrase- or word-medially to the neighboring voiced environment with voicing during the closure. As Abramson (pers. comm.) pointed out, this is interesting because it is not the fortis stops with the shortest VOT of the three categories but rather the lenis stops with the intermediate VOT that are voiced in intervocalic positions (Lisker & Abramson, 1964). This suggests that these stops differ in many dimensions, including the tension of the vocal folds and aerodynamic properties, and that VOT is not the major feature distinguishing the stops.

Han & Weitzman (1970) noted that, in addition to the VOT differences, different acoustic features can be observed in the onset phase of voicing following the stop release. They reported that $f0$ and the intensity characteristics of the vowel onset also contribute to the manner distinction. In general, $f0$ after aspirated and fortis stops is relatively higher than after lenis stops (cf. Hardcastle, 1973; Kagaya, 1974; Jun, 1993; M.-R.C. Kim, 1994; Cho, 1996). These studies indicate that $f0$ contrasts serve as a supplementary cue to distinguish lenis stops from fortis and aspirated stops, but not necessarily to distinguish aspirated from fortis stops, a finding that was later supported by a perception test (J.-I. Han, 1996). Furthermore, Jun (1993, 1996a) found that the $f0$ difference is not simply due to the phonetic perturbation of the preceding consonant that has been noticed in other languages (Lehiste & Peterson, 1961; Lieberman, 1963; Ladefoged, 1964; Hombert, 1978; Hombert, Ohala & Ewan, 1979, *inter alia*). In Korean, it is phonologically encoded in the intonation system in most dialects of Korean (including the Seoul dialect), so that a phrase-initial syllable beginning with a fortis or an aspirated stop is realized with a high (H) tone and that with a lenis stop is realized with a low (L) tone.

Han & Weitzman (1970) reported that the harmonic components are weak for lenis stops, intermediate for aspirated stops, and strong for fortis stops. They further argued that these observations are indicative of a difference in the intensity buildup following the voice onset associated with each stop: relatively more time is needed for glottal intensity to build up following a lenis stop or an aspirated stop than that following a fortis stop.

The duration of the stop closure has been considered as another acoustically distinctive feature associated with Korean stops—in general, the stop closure is shortest for

lenis stops, intermediate for aspirated stops, and longest for fortis stops (Silva, 1992; M.-R.C. Kim, 1994; J.-I. Han, 1996). But an electropalatographic (EPG) study by Cho & Keating (2001) showed that there is no significant difference between aspirated and fortis stop closure durations, all else being equal. Cho & Keating (2001) also examined linguopalatal contact (the contact between the tongue and the roof of the mouth) for different stops using EPG and found that lenis stops have less linguopalatal contact than aspirated or fortis ones (see also Cho, 1998). The longer duration and the greater linguopalatal contact associated with both the aspirated and the fortis stops indicate that they are articulatorily "strong" stops, compared to the lenis stop.

The nature of the aspirated and the fortis stops has also been examined with reference to laryngeal and supralaryngeal articulatory tension. One of the earlier studies, C.-W. Kim (1965), characterized the stops in terms of two features: "tension" of the articulation, which served to distinguish aspirated and fortis stops from lenis stops; and "aspiration", which served to distinguish aspirated stops from fortis stops (and lenis stops). Some of the supportive findings for "tension" associated with fortis and aspirated stops were a faster rate of vocal fold vibration after release, greater amplitude of pressure, longer duration of increased pressure, faster pressure buildup, greater linguopalatal contact, and greater lip muscle activity (for bilabial stops). In line with C.-W. Kim, Hardcastle (1973, p. 271) suggested on the basis of measurements of VOT, frequency of glottal cycle at vowel onset and airflow rate that "a feature 'tensity', defined in terms of isometric muscular tension in the vocal cords [i.e., increased muscular tension of both agonist and antagonist muscles so that there is no change in the length of the vocal folds], can usefully be employed to explain some of these properties".

Articulatory aspects of Korean stops have become better understood by examination of the glottal configuration. C.-W. Kim (1970), using cineradiographic evidence, reported that the glottal opening is larger for aspirated stops, intermediate for lenis stops, and narrower for fortis stops, arguing that the degree of aspiration is proportionally corre-lated with the degree of glottal opening at the time of release of the oral closure. Kagaya (1974), in a fiberscopic study, found that fortis stops have approximated vocal folds well before the articulatory release, while the glottis is quite open for the lenis stop at the time of the release, but not as large as for the aspirated stop. Kagaya suggested that both the fortis and aspirated stops are characterized by some intrinsic laryngeal gestures. In his view, fortis stops can be characterized by (1) a completely adducted state of the vocal folds before the articulatory release, (2) stiffening of the vocal folds, (3) abrupt closure of the vibrating vocal folds near the voice onset, (4) increasing subglottal pressure, and (5) lowering of the glottis immediately before the release. Aspirated stops are associated with positive abduction of the vocal folds and heightened subglottal pressure. On the other hand, none of these positive laryngeal gestures are observed for lenis stops. More recently, Jun, Beckman & Lee (1998), in their fiberscopic study of vowel devoicing, found a similar result in terms of the timing and size of the glottal opening for the three stop types.

The glottal state in these stops has also been examined by electromyography (EMG), which allows the investigation of the activities of the intrinsic laryngeal muscles. Hirose *et al.* (1974) reported that fortis stops are characterized by a sharp increase in thyro-arytenoid activity before the stop release, which presumably resulted in increased tension of the vocal folds and constriction of the glottis during or immediately after the stop closure. In aspirated stops, all activity of the adductor muscles of the larynx is suppressed immediately after the articulatory release. A steep increase in activity of the adductor

muscles always followed this suppression, presumably due to the movement into the position for voicing. In lenis stops, the suppression of adductor muscle activity is not significantly involved as compared with aspirated stops, and there is no transient increase in thyroarytenoid activity before the articulatory closure. These results suggest that a simple dimension of adduction–abduction of the vocal folds is not sufficient in characterizing Korean stops, as implied in the studies of C.-W. Kim (1970) and Kagaya (1974).

Dart (1987) investigated the different aerodynamic properties of fortis and lenis stops in Korean. She measured intraoral air pressure and oral flow associated with the fortis and lenis stops in prevocalic position. One of the main results in her study is that the production of fortis stops is characterized by "a higher intraoral pressure before release, yet a lower oral flow after release", which is somewhat counterintuitive since in general higher intraoral pressure is associated with greater oral flow. Dart's aerodynamic modeling accounted for the pressure-flow relations by modeling fortis stops with greater vocal tract wall tension and tightly adducted vocal folds before the articulatory release.

As it became clear that the different types of stops are associated with different glottal configurations in the production of not only the stops themselves but also the onset of the following vowel, researchers began to raise the question of whether the voice quality of the vowel is influenced by the preceding consonant. They hypothesized that the voice quality of the vowel is similar to a breathy voice after the lenis stop and to a laryngealized or "pressed" voice after the fortis stop. A laryngographic study by Abberton (1972) showed that the onset of vowels after fortis stops has some of the characteristics of creaky voice with a long closed phase and a slow opening phase. N. Han (1998) reported that vowels after lenis stops have a breathy voice as indicated by positive $H1-H2$ values (the difference in amplitude between the first and the second harmonics), but she found that vowels after fortis stops do not always have negative $H1-H2$ values, which, if present, would have indicated a laryngealized or pressed voice quality.

1.1.2. *Fricatives*

Korean has two denti-alveolar fricatives. As in the stops, there is no voicing contrast. But, unlike the stops, there is only a two-way contrast between /s/ and /s*/. As mentioned earlier, /s*/ has been called a tense or fortis fricative, but the categorization of /s/ has been controversial. In Korean orthography, the plain /s/ is regarded as lenis and behaves that way in phonological processes. But its behavior in phonetic processes and its phonetic realizations are generally believed to be similar to stops in the aspirated category. For example, /s/ becomes tense after a lenis stop (e.g., /paksa/ → [paks*a] "Ph.D") as does the lenis stop (e.g., /pakta/ → [pakt*a] " to pin down"). On the other hand, it is also similar to the aspirated stop in that it is not likely to become voiced between voiced segments, and it generally triggers a high tone in the beginning of an Accentual Phrase (an intonationally defined minor phrase smaller than an Intonation Phrase (Jun, 1993)). Fiberscopic data in Kagaya (1974) and Jun *et al.* (1998) showed that /s/ has a glottal opening configuration similar to aspirated stops, and Jun *et al.* further showed that /s/ has a larger glottal opening than /s*/. Based on acoustic data, Yoon (1998) and Park (1999) also showed that /s/ has some degree of aspiration as in aspirated stops, and Park (1999) claimed, based on the energy difference between the first and second harmonics at vowel onset, that the plain /s/ should be categorized in the aspirated category because the vowel onset after plain /s/ is breathier than after fortis /s*/. However, the relative

breathiness associated with /s/ cannot be a direct metric for whether /s/ has the characteristics of an aspirated segment or not, since the lenis segment, as noted earlier, can also be associated with breathy voicing in the following vowel onset. Furthermore, an EPG study by S. Kim (2001) showed that /s/ has a significantly less linguopalatal contact with a wider frication channel than /s*/, which is a result similar to the difference in the amount of linguopalatal contact between the lenis and the fortis stops.

2. Methods

2.1. *Data acquisition and speakers*

2.1.1. *Speakers*

The Cheju data are based on recordings of eight male speakers of Cheju, in their mid 50s and mid 70s. Three of them lived in Cheju city and the other five lived in the mountainous area of Southern Cheju. The Seoul data were collected from four male speakers of Seoul in Los Angeles, California, in their late 50s and early 60s. Two of them were short-term visitors to America, and two others moved to America 10–13 years ago, but did not speak English as they were living in communities in which virtually only Korean was spoken. All of the speakers were literate in Korean.

2.1.2. *Data acquisition procedures*

Each of the speakers was recorded using a close-talking, noise-canceling Shure microphone and a Sony DAT recorder. For the Cheju data, the recordings were made in the home of one of the speakers, or in a quiet part of the village community center. For the Seoul data, two speakers (visiting scholars) were recorded in the UCLA phonetics lab, and the other two speakers were recorded in their home or office.

For the Cheju speakers, each word was written on an index card in Korean orthography reflecting the sounds of Cheju as closely as possible. To help the speakers produce each word more naturally, a word triggering the context was written next to the target word on the index card. For example, a word for "baby" was written for the target word "to give birth"; a word for "ground" was written for the target word "to dig". Speakers produced each word twice after being prompted by one of the authors, Jun, or by a Cheju-native linguist (see Acknowledgements). The word list was rehearsed with each speaker before the recording was made. Data collection procedures were similar for Seoul speakers except that each word was written in the standard orthography.

In this study, acoustic analysis techniques were used to investigate properties of all the obstruents falling within the scope of the study, and aerodynamic techniques were used to investigate the properties of a subset of obstruents, namely the bilabial stops. For the aerodynamic data, oral airflow and pressure were recorded using the Macquirer X16 system. Speakers held a face mask against the lower part of the face, below the nose, capturing all the oral airflow. They also held a tube (internal diameter 2 mm) between their lips to record the pressure of the air in the mouth. A microphone within the face mask recorded the audio signal. Speakers found little difficulty in producing the required phrases in these conditions, and the audio signal indicated that the utterances sounded reasonably natural. Only bilabial stops, as shown in Table V, can be investigated in this way. (See Ladefoged (1997) for detailed procedure.) The speakers were too elderly and

unaccustomed to phonetic experimentation to ask them to pass a tube through the nose so that pressures behind dental and velar closures could be recorded. Recorded materials were digitized at a sampling rate of 22 050 Hz and analyzed using the Kay Elemetrics's Computerized Speech Lab (CSL). The flow and pressure signals were sampled at a rate of 2 kHz and the audio signal was sampled at 10 kHz. They were analyzed using SciconRD's Macquirer.

2.2. *Speech material*

For the acoustic study, all nine stops of Korean were examined. Each stop was placed in the initial position of a word in isolation, and was followed by the open vowel /a/. Table III shows the word list in a phonemic transcription. Note that for the Cheju lenis /t/, two different words were used because some speakers were more familiar with one than the other.

The fricatives /s, s*/ were recorded in word-initial and medial positions, in the words shown in Table IV. The approximant /l/ in word initial position (i.e., /latio/ "radio") was included to provide a baseline when comparing the voice qualities of vowels conditioned by the type of preceding consonant (see Section 3.4). Each word in Tables III and IV was read six times by each Seoul speaker. The Cheju speakers repeated each word only twice, because of the need to collect data from as many speakers as possible in the limited time period available for fieldwork.

TABLE III. Stops recorded for acoustic measurements

		Seoul		Cheju
pʰ	pʰata	"to dig"	pʰamtʃə	"to dig"
p*	p*ata	"to grind"	p*amtʃə	"to squeeze oil"
p	pata	"sea"	pataŋ	"sea"
tʰ	tʰata	"to ride"	tʰamtʃə	"to get (shy)"
t*	t*ata	"to pick"	t*api	"ground"
t	talta	"to be sweet"	talkwatʃəmtʃə	"to heat iron"
			tɐlanamtʃə	"to run away"
kʰ	kʰatɨ	"card"	kʰamtʃə	"to get burned"
k*	k*ata	"to peel"	k*amtʃə	"to peel"
k	kata	"to go"	kamtʃə	"to go"

TABLE IV. Fricatives recorded for acoustic measurements

		Seoul		Cheju
Word-initial				
s	sata	"to buy	salamtʃə	"to live"
s*	s*ata	"to wrap"	s*amtʃə	"to wrap"
Word-medial				
s	nasata	"(this) is a bolt"	asakatʃə	"to take away"
s*	pas*ak	"completely (dried)"	pʰas*ak	"breaking noise"

TABLE V. Word list for aerodynamic study

		Seoul		Cheju
p^h	p^{h}e	"card"	p^{h}e	"card"
p	pe	"ship"	pe	"ship"
p*	p*e	a stem of "to draw"	p*e	"bone"

For the aerodynamic study, a corpus with only bilabial stops in word-initial position was included in the current study as shown in Table V. The items in Table V were read nine times each by each Seoul speaker, and four times each by each Cheju speaker (except for one speaker in the rural area who did not wish to participate in the aerodynamic study, but did participate in the rest of the study).

2.3. *Measurements*

In this section, we will describe the measurement criteria employed to investigate the acoustic and aerodynamic properties of stops and fricatives. For stops, we will consider VOT, burst energy, fundamental frequency ($f0$), amplitude difference between the first harmonic and the second harmonic ($H1-H2$), amplitude difference between the first harmonic and the second formant ($H1-F2$), intraoral pressure and airflow data. For fricatives, we will consider the duration, centroid of fricative noise, RMS energy, $f0$, $H1-H2$, and $H1-F2$, with particular reference to the contrasting nature of /s/ and /s*/.

2.3.1. *Measurements for stops*

Voice onset time. VOTs for all stops were taken from the point of the stop release to the voice onset of the second formant in the following vowel, as seen in spectrograms. Thus, for the lenis stops, any breathy voicing with only low-frequency harmonics was included in the VOT.

Relative burst energy. The acoustic energy at the burst and in the middle of the following vowel were measured from an acoustic energy profile, using a 10 ms window. To normalize the energy difference across speakers, the percentage value of the burst energy relative to the energy at the midpoint of the vowel was employed. Greater burst energy for a stop can be expected in two cases—when a consonant (i.e., /t/) has a relatively smaller amount of linguopalatal contact, resulting in a fast release, as opposed to a larger contact area (cf. Stevens, Keyser & Kawasaki, 1986), and when the airflow is greater at the release (which is presumably due to a greater air pressure behind the constriction immediately before the release).

Fundamental frequency ($f0$). $f0$ was measured at the onset and the midpoint of the vowel, using the pitch track along with the first harmonic values from an FFT with a 25 ms window as supplementary checks. From the $f0$ differences, we can infer some physical information about the vocal folds (e.g., tension or stiffness) associated with different consonant types.

$H1-H2$ and $H1-F2$. Energy values (dB) for the first ($H1$) and second ($H2$) harmonics and the peak harmonic forming the second formant ($F2$) were taken at the onset and midpoint of the vowel, using FFT spectra with a 25 ms window (40 Hz bandwidth). In

addition, harmonic values were also measured at the corresponding points of the vowel after the liquid /l/, which was used as control data representing modal voicing. To avoid the influence of the first formant energy on the amplitude of $H1$ and $H2$, low vowels were used for the target words.

The difference in amplitude between $H1$ and $H2$ has been frequently used to distinguish between breathy and modal voicing (e.g., Bickley, 1982; Ladefoged, 1983; Huffman, 1987; Klatt & Klatt, 1990; Blankenship, 1997, 2002). Breathy voicing is produced with a relatively larger open quotient with the vocal folds remaining closed for a shorter time (e.g., open for 80–100% of the cycle for breathy phonation, *vs.* 65–70% for modal phonation (Childers & Lee, 1991)). As a result, the spectrum is dominated by the energy at the fundamental frequency, resulting in the amplitude of $H1$ being markedly higher than that of other harmonics. On the other hand, laryngealized vowels with creaky voicing (or "pressed voicing" in Stevens' (1999) term) would have the opposite result, as they are produced with a smaller open quotient with the vocal folds remaining closed for a longer time. Thus, a greater $H1-H2$ would indicate breathiness of the vowel and a smaller or negative $H1-H2$, would indicate "pressed" voicing quality.

$H1-F2$. The spectral slope, as obtained from $H1-F2$, was measured at the onset and the midpoint of the vowel. This is an indicator of the abruptness of vocal fold closure. A gradual adduction of the vocal folds mainly excites the lower resonances of the vocal tract and, as a result, the sound wave is nearly sinusoidal with most energy near $f0$. The resulting spectrum has a steep downward slope from which a greater $H1-F2$ is expected. We expect that any breathy voicing should be associated with a gradual adduction with a greater $H1-F2$. On the other hand, when there is an abrupt adduction with the vocal folds coming together all at once, which is one of the characteristics of pressed or creaky voicing, the abruptness of the closure excites a wider range of frequencies. As a result, the sound wave has a spectrum with energy spread across a higher range of frequencies. Schematized spectra for different types of voicing are shown in Fig. 1, based on Stevens (1999, pp. 86, 90).
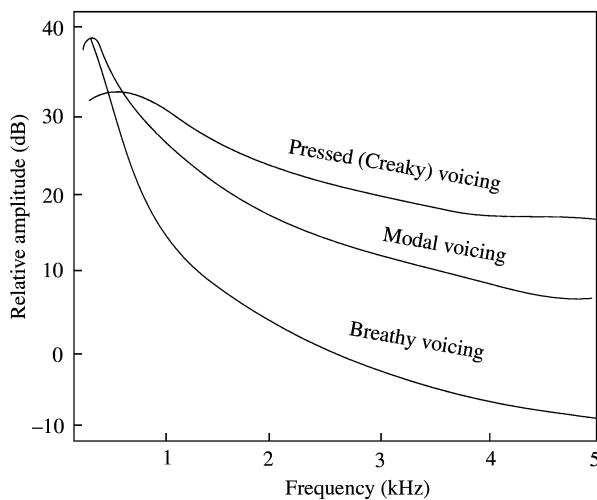
Figure 1. Schematized spectra for different phonation types, based on Stevens (1999, pp. 86, 90).
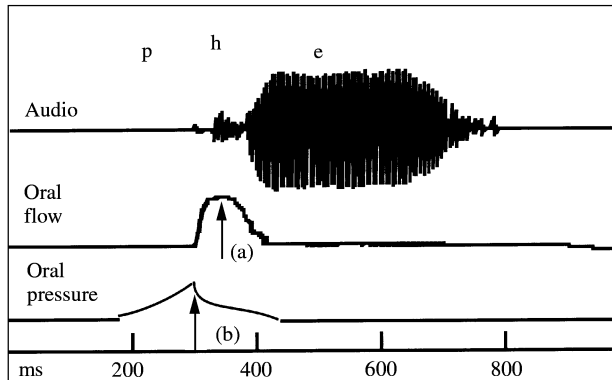
Figure 2. Waveforms of audio, oral airflow and oral pressure. The points marked
by arrows indicate peak oral flow (a) and oral pressure (b).

Oral pressure and flow. The maximum flow after the release of the closure and the peak
oral pressure during the closure were measured, as indicated by arrows (a) and (b) in
Fig. 2, respectively. Oral pressure and flow provide information about the degree of
glottal constriction (cf. Ladefoged, Maddieson & Jackson, 1987). Similarly, as we have
noted, Dart (1987) using these measures found indications of differences in the glottal
configuration immediately after the release of the fortis and lenis Korean stops.

2.3.2. *Measurements for fricatives*

Fricative duration. Duration measurements were made from the spectrograms of each
token for /s/ and /s*/ both in word-initial and word-medial positions. For /s/, we also
measured the duration of both frication and aspiration, separately and combined.

Centroid of the fricative noise. The centroid is the center of gravity of a defined part of
the spectrum, each frequency being weighted according to its amplitude (cf. Forrest,
Weismer, Milenkovic & Dougall, 1988; Zsiga, 1993). Centroid values were taken from
FFT spectra in the frequency range from 500 to 10,000 Hz, using a 25 ms window
centered around the midpoint of the fricative portion. For denti-alveolar fricatives, /s/
and /s*/, a high centroid is expected as the source is filtered by the front cavity reson-
ance, resulting in a spectrum peak in the vicinity of $F4$ or $F5$ (Stevens, 1999). We were
particularly interested in determining whether there was a difference in the centroid
frequency of /s/ and /s*/, which would indicate the difference in the size of the front
cavity—the higher the frequency, the smaller the front cavity.

$H1 - H2$ *and* $H1 - F2$. As measured for stops, $H1 - H2$ and $H1 - F2$ were measured at
the onset and the midpoint of the vowels that follow /s/ and /s*/.

Fundamental frequency ($f0$). As measured for stops, $f0$ was taken at the onset and the
midpoint of the following vowel.

2.4. *Data analysis*

The data were averaged over repetitions so that each subject contributes data to only one
experimental unit. This reduction in data was done in order to meet the statistical
assumption, known as "the assumption of independence of the error effects" (Kirk, 1995)
which assumes that, in statistical analyses such as Analysis of Variance, each data point
that is entered should correspond to the score for an individual subject. In other words,

this data reduction procedure was done to avoid overestimation of the effect (Type I or *alpha* error).

Evaluation of the systematic influence of various factors was made based on repeated measures General Linear Model (GLM) Analyses of Variance (ANOVAs). The within-subject factor considered in most cases was Stop Category (Lenis, Fortis, Aspirated) (or Fricative Category (s, s*) for fricatives). Other within-subject factors that were considered include Place of Articulation (Bilabial, Alveolar, Velar) for VOT, and Position-in-Vowel (onset, midpoint) for $f0$, $H1 - H2$, $H1 - F2$. In addition, since we were also interested in whether Seoul and Cheju dialects would differ in phonetic characteristics of stops and fricatives, the Dialect (Seoul, Cheju) was considered as a between-subject factor in a mixed model of repeated measures ANOVAs (with at least one within- and one between-subject factors). For the same reason, we will also present graphical results in figures separately for each dialect regardless of whether or not there is an effect of Dialect or an interaction between Dialect and other factors.

In conducting repeated measures ANOVAs, data were also adjusted to meet another statistical assumption, namely, the sphericity assumption that the variance of the difference scores for all pairs of treatment levels are homogeneous (Huynh & Feldt, 1970). In order to avoid violation of the sphericity assumption, a Huyhn–Feldt corrected degree of freedom was used in generating F ratio and p values. (See Max & Onghena (1999) for a tutorial on statistical issues of reduction to experimental units and the sphericity assumption.) In reporting the results of repeated measures ANOVAs, p-values less than 0.05 will be considered significant. In addition, p-values less than or equal to 0.08 will be reported as non-significant trends (*ns trend*) in order to make note of some tendencies that might otherwise be neglected due to the adjustment of degrees of freedom or data reduction.

In order to further analyze the observed data within a factor, *post hoc* analyses were performed. In case the data have heterogeneous variances and the observations on a given subject are correlated in the cases of repeated observations, the Bonferroni/Dunn model (cf. Hays, 1994) was used, but at the lowered alpha level of 3% (equivalent to $p < 0.01$ when there are three levels to compare within a factor.)

Finally, whenever there was an interaction between Dialect and other factors, we consulted mean values for each speaker, so that we could see whether any observed patterns were homogeneously found for speakers within a dialect. In addition, we employed cluster analysis with a hierarchical clustering technique that starts out with each point being its own cluster, then at each step combining the clusters that are close to each other. Such a cluster analysis with Speakers as a grouping variable allowed us to examine as to how speakers are grouped together regardless of dialects and whether any observed dialectal difference in a dependent variable is in fact attributable to speaker variation.

The SPSS statistical package (version 10, SPSS Inc., 1999) and StatView (version 5.0.1, SAS Inc., 1999) were used to perform repeated measures ANOVAs and *post hoc* comparisons, respectively.

3. Results and discussion for stops

3.1. *VOT*

Results of repeated measures ANOVA (Stop Category by Place of Articulation by Dialect) showed main effects of two within-subject factors Stop Category ($F(1.89,$
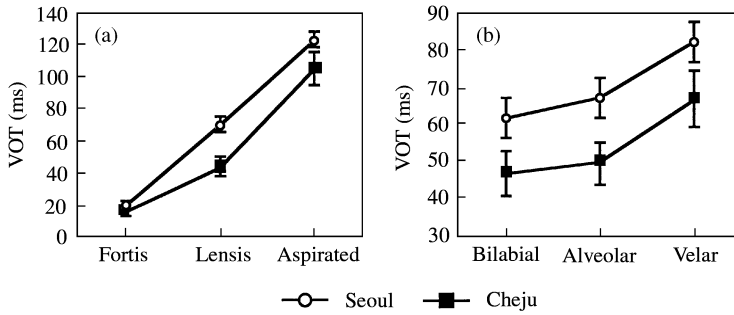
Figure 3. Variation of VOT: (a) as a function of stop category and (b) as a function of place of articulation. Error bars indicate standard errors.

18.68) $= 141.61$, $p < 0.0001$) and Place of Articulation ($F(2, 10) = 34.89$, $p < 0.0001$) on VOT. However, there was only a non-significant trend for Dialect ($F(1, 10) = 3.89$, $p = 0.078$), and no interaction between factors.

The results are summarized in Fig. 3. First, pairwise Bonferroni/Dunn *post hoc* comparison showed that VOT is shortest for the fortis stop, intermediate for the lenis stop, and longest for the aspirated stop ($p < 0.0001$ for each comparison), which is true for both Seoul and Cheju Korean, as seen in Fig. 3(a). Second, there is a potential dialectal difference in VOT as an *ns* trend—i.e., VOT for Seoul Korean is in general longer than that for Cheju Korean.

It is interesting to compare this potential dialectal difference with diachronic differences observed in the literature (e.g., Silva, 1992; N. Han, 1998). For example, reviewing the literature about Korean VOTs over the past 50 years or so, N. Han observed that VOT for the aspirated and the lenis stop have increased over time. We have reported some evidence elsewhere (Cho *et al.*, 2001) that Cheju speakers are in general conservative in terms of preserving vowel sounds. It is then conceivable that the overall shorter VOT in Cheju is due to the conservatism of the Cheju people who are isolated from mainland Korea. This potential dialectal difference is even clearer for the lenis stop—i.e., the mean VOT for the lenis stop in Cheju (47 ms) is about two-thirds of that in Seoul (72 ms). Each Seoul speaker's mean VOT for the lenis stop is always longer than any Cheju speaker's mean VOT except for one Cheju speaker (C5) who has a mean of 73 ms. In the cluster analysis, this Cheju speaker is grouped together with the four Seoul speakers, suggesting that he may be influenced by the standard Seoul dialect. Nonetheless, all the Seoul speakers are clustered together, implying that at least Seoul speakers all behave similarly with respect to VOT, having relatively longer VOT than a majority of Cheju speakers.

Finally, regarding the effect of place of articulation, the VOT of the velar stop is significantly longer than that of either the bilabial or the denti-alveolar stops ($p < 0.0001$), as seen in Fig. 3(b). There is no significant difference between the bilabial and the denti-alveolar VOTs. This pattern remains true for both Seoul and Cheju Korean with no interaction between Place of Articulation and Dialect. This is, in general, in agreement with the tendencies noted in other languages (Cho & Ladefoged, 1999).

3.2. *Relative burst energy*

Results of repeated measures ANOVAS (Stop Category by Place of Articulation by Dialect) show that there is a significant effect of Stop Category on the relative burst
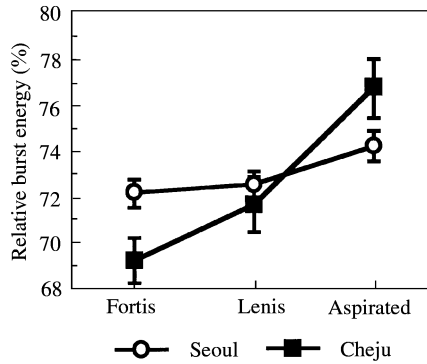
Figure 4. The relative burst energy (%) for different stop categories. Error bars indicate standard errors.

energy ($F(2, 20) = 8.44$, $p < 0.01$) but a weak effect of Place of Articulation as a non-significant trend ($F(2, 20) = 3.94$, $p = 0.69$). There is neither an effect of Dialect nor any interaction between Dialect and other factors.

Bonferroni/Dunn *post hoc* comparisons revealed that the aspirated stop has significantly greater burst energy than the other two types of stops ($p < 0.0001$) as shown in Fig. 4, while there is no difference between the fortis and the lenis stops. However, in Fig. 4, for Cheju Korean we can see that the average relative burst energy (%) is clearly greater for lenis than for fortis stops, which is not the case for Seoul Korean. While this seems to be a potential dialectal difference, the pattern of lenis > fortis is in fact observed for only two out of four Seoul speakers and five out of eight Cheju speakers, suggesting that the observed difference in the figure is due to speaker variation regardless of Dialect, which is why there is no significant Dialect by Stop Category interaction.

There are two explanations of the greater burst energy associated with stops: (a) a smaller linguopalatal contact (Stevens *et al.*, 1986), which might be associated with a greater speed of movement, and (b) a greater intraoral airflow at the release. However, Cho & Keating (2001) found that there is no difference between the fortis and aspirated stops in the extent of the linguopalatal contact. Both of them have significantly greater contact than the lenis stops. This indicates that the amount of the contact itself is not the primary source for the difference in burst energy. The aerodynamic factor, the intraoral airflow, provides a more likely explanation. Indeed, our aerodynamic results reported in Section 3.5.2, show that the aspirated stop is produced with the greatest airflow. This is consistent with it having the greatest burst energy.

3.3. *f0 differences*

A three-way repeated measures ANOVA (Stop Category by Position-in-Vowel (onset *vs.* midpoint) by Dialect) showed that $f0$ was significantly influenced by Stop Category ($F(2, 20) = 38.689$, $p < 0.0001$), but not by Position-in-Vowel. A non-significant trend was found for the effect of Dialect ($F(1, 10) = 3.99$, $p = 0.072$). As for the effect of Stop Category, Bonferroni/Dunn pairwise *post hoc* comparisons showed that $f0$ for the lenis stop was lower than the other two stops at $p < 0.001$. This pattern can be seen in Fig. 5. In the figure we can also see that $f0$ is significantly higher for Cheju Korean than for
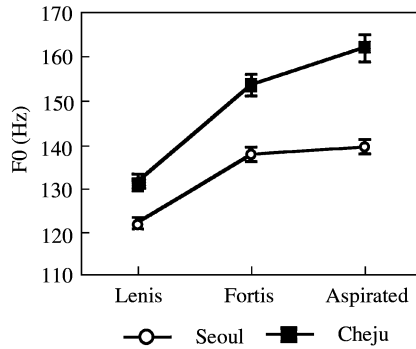
Figure 5. $f0$ differences in the following vowels. Data are pooled across position-in-vowel (see text). Error bars refer to standard errors.

Seoul Korean at $p < 0.01$ (Bonferroni/Dunn). However, cluster analysis conducted separately for each stop category showed that speakers are not clearly grouped into two dialects, suggesting that the observed non-significant trend effect of Dialect may be simply due to speaker variation.

In addition, there was a significant interaction between Stop Category and Position-in-Vowel ($F(2, 20) = 5.539$, $p < 0.05$). This interaction is primarily due to the fact that $f0$ increases from the onset of the vowel to the midpoint for lenis stops, while no such difference between the onset and the midpoint of the vowel is associated with aspirated and fortis stops. No other interactions were found between factors.

3.4. $H1-H2$ and $H1-F2$

A repeated measures ANOVA (Stop Category by Position-in-Vowel by Dialect) indicates a main effect of Stop Category on $H1-H2$ ($F(2, 20) = 62.49$, $p < 0.0001$). There is a non-significant trend for Position-in-Vowel ($F(1, 10) = 4.71$, $p = 0.055$) but no effect of Dialect. However, there are significant interactions of Stop Category by Position-in-Vowel ($F(2, 20) = 15.65$, $p < 0.001$), and Stop Category by Position-in-Vowel by Dialect ($F(2, 20) = 3.97$, $p < 0.05$).

Pairwise *post hoc* comparisons show that stops differ significantly from each other at $p < 0.001$, which is true for both dialects—$H1-H2$ is greatest (positive) for lenis stops, intermediate for aspirated stops, and smallest (negative) for fortis stops, as shown in Fig. 6. This suggests that the onset of the vowel (the first 30 ms period) has more breathy voicing immediately after the lenis stop, and more pressed (or laryngealized) voicing after the fortis stop. The voice quality of the vowel after the aspirated stop is close to the modal voice: its $H1-H2$ value is about the same as the control (modal) value taken from the onset of the vowel after /l/, which has no known effect on the phonation type of the following vowel. This is consistent with Hirose *et al.*'s (1974) observation that aspirated stops are produced with a steep increase in the activity of the adductor muscles of the larynx following the suppression of the adductor muscles immediately after the articulatory release, while lenis stops do not show such robust adductor muscle activity. This appears to contribute to the fact that aspirated stops are not associated with breathy voice even though they have the largest glottal opening during the closure. What is especially interesting about the overall finding is that both the fortis and the lenis stops
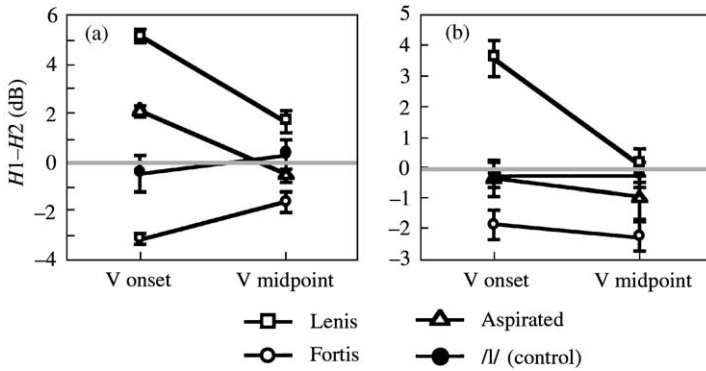
Figure 6. Difference between the amplitudes (dB) of the first harmonic ($H1$) and the second harmonic ($H2$) for different stop categories and /l/. Error bars indicate standard errors. (a) Seoul and (b) Cheju.
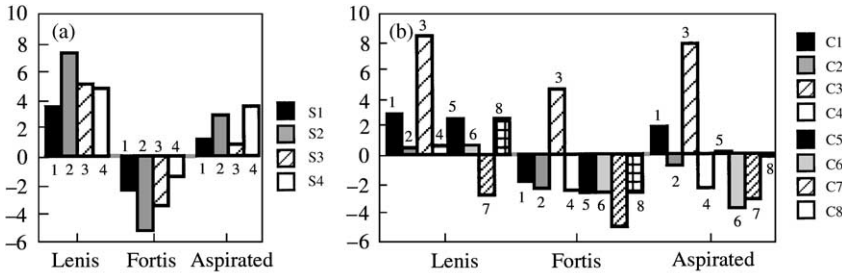
Figure 7. Difference between the amplitudes (dB) of the first harmonic ($H1$) and the second harmonic ($H2$) separated by each speaker. The numbers in the figures identify the different speakers. (a) Seoul and (b) Cheju.

appear to be associated with nonmodal phonation, departing from modal phonation in opposite directions. These differences are significantly maintained in the midpoint of the vowel consistently across dialects, though the difference becomes smaller at the mid-point.

Considering the dialect/speaker difference, it is especially worth examining speakers separately since paralinguistic differences in voice quality among speakers may have caused the results described in the preceding paragraph. Fig. 7 shows the $H1-H2$ difference for each speaker separately. A rather obvious point is that the absolute values of $H1-H2$ vary with speakers, showing an overall speaker variation in voice quality. Nonetheless, what is consistent across speakers in both dialects is that the lenis stop is associated with a larger $H1-H2$, and the fortis stop, with a smaller $H1-H2$. This leads us to generalize with confidence that speakers linguistically control the voice quality in order to differentiate at least the lenis and the fortis stops. However, what is inconsistent across speakers or perhaps across dialects is whether or not $H1-H2$ values for the aspirated stop fall in between those for the lenis and the fortis stops. All four Seoul speakers make a clear three-way distinction among stops, showing a pattern of fortis < aspirated < lenis in $H1-H2$. Note that this is inconsistent with the findings of

Ahn (1999), who examined differences in the voice quality of the vowels as a function of the preceding stop category. He found that both raw and normalized $H1-H2$ values for the fortis stop were smaller than those for the lenis stop, whereas those for aspirated stops were greater, giving an order of fortis < lenis < aspirated. We do not know as to why there is a difference between Ahn's and our findings. While the difference could be due to the relatively younger age group of Ahn's speakers (on average 36.5 years old), it may also be possible that the procedural difference may have caused the inconsistency between the two studies.

In Cheju Korean, $H1-H2$ of the aspirated stop is close to that of the lenis stop for some speakers (e.g., C1, C3, C7), or to that of the fortis stop for some others (e.g., C4, C6). The other three speakers (C2, C5 and C8) make a clear three-way distinction, as do Seoul speakers. This suggests that the voice quality associated with the aspirated stop does not matter too much, perhaps because the other phonetic cues (e.g., large aspiration) may be enough to differentiate the aspirated stop from the other two stop categories, or perhaps because the language simply chose not to rely on $H1-H2$ for identifying aspirated stops. Considering this within-dialect variation in Cheju Korean, the only veridical generalization to be made about dialect/speaker differences is that Seoul speakers consistently make a relatively clear three-way distinction (fortis < aspirated < lenis) in $H1-H2$ while only some of Cheju speakers do, which suggests that the seemingly dialectal difference may well be ascribable simply to speaker differences irrespective of the dialectal factor.

We now turn to the spectral slope ($H1-F2$), a hypothesized indicator of the abruptness of vocal fold closure. A repeated measures ANOVA indicated a significant effect of Stop Category on $H1-F2$ ($F(1.29, 17.09) = 17.753$, $p < 0.0001$), but there was no main effect of Position-in-Vowel and Dialect. No significant interactions between factors were found.

Post hoc comparisons show that the fortis stop has the smallest (negative) $H1-F2$ while the lenis and aspirated stops have positive $H1-F2$s, as shown in Fig. 8. This direction is consistently true cross-dialectally for both the onset and the midpoint. This indicates that the closing phase of the vocal folds during the voicing is more abrupt for the fortis stop than for the other stops, even in the middle of the following vowel. The vowel is produced with the folds coming together more rapidly after the fortis stop, providing more energy in the second or higher formant region, presumably because they are under greater tension.
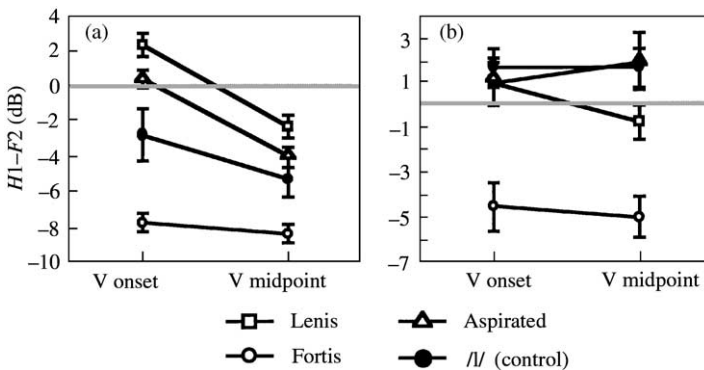
Figure 8. Difference between the amplitudes (dB) of the first harmonic ($H1$) and the second formant ($F2$). Error bars indicate standard errors. (a) Seoul and (b) Cheju.

To sum up, what emerges from the $H1-H2$ and the $H1-F2$ data is that the stops can be further differentiated at least in part by the voice quality of the following vowel. Despite some potential dialectal and speaker differences, the common result is that vowels after the fortis stop are relatively pressed with accompanying abruptness of the adduction gesture, but vowels after the lenis stop are breathy with a relatively less abrupt adduction gesture.

3.5. *Aerodynamics*

3.5.1. *Intraoral air pressure* (*Po*)

We will first examine the results of intraoral air pressure (Po). A repeated measures ANOVA (Stop Category by Dialect) suggests that intraoral air pressure is significantly influenced by Stop Category ($F(1.81, 18.11) = 6.78$, $p < 0.005$), but not by Dialect. As shown in Fig. 9(a), the peak oral pressure during the closure in both dialects is smaller for the lenis stop than for the fortis and aspirated stops at $p < 0.01$, while there is no difference between the latter two. There is no significant Stop Category by Dialect interaction.

The oral pressure during a stop closure is dependent on four factors: First, the subglottal pressure, which is its ultimate source; second, the glottal impedance, which results in the oral pressure being less than the subglottal pressure; third, the tension of the walls of the vocal tract, which, if allowed to expand, will result in a lowering of the oral pressure; and fourth, the duration of the stop closure, which, in the case of initial stops when the subglottal pressure is still increasing, would result in stops with a long closure having a higher oral pressure. From the oral pressure data alone, it is difficult to assess as to which of these factors is most responsible for the higher pressure in the aspirated and fortis stops in comparison with the lenis stops.

Dart's model suggested that the first factor, an increase in the subglottal pressure due to an increase in respiratory effort, was the most likely cause of the higher oral pressure in fortis stops. As also noted by Dart, we should note that the second factor, the glottal impedance, might lower the oral pressure in fortis stops in comparison with aspirated stop, as the fortis stop is produced with constricted or approximated vocal folds well before the release (C.-W. Kim, 1970; Kagaya, 1974; Jun *et al.*, 1998). But insofar as the fortis stops have oral pressures that are similar to those in aspirated stops, it seems that
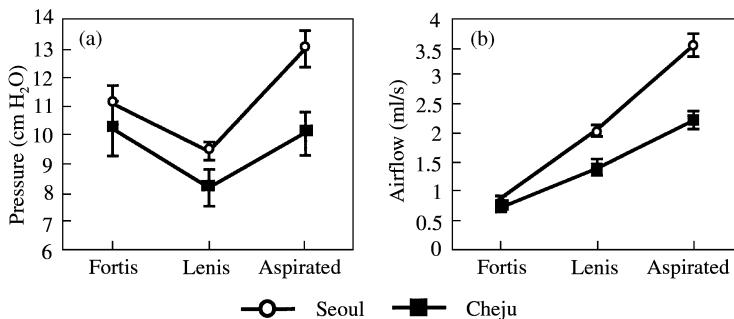
Figure 9. (a) Oral pressure (cm H_2O) and (b) airflow (ml/s). Error bars indicate standard errors.

the fortis stops may have an even higher subglottal pressure than the aspirated stops.

The third factor, the tension of the walls of the vocal tract, may also account for the higher Po for the fortis stop, which is arguably produced with a relatively more tense vocal tract wall than the lenis stop (Kagaya, 1974; Dart, 1987). C.-W. Kim (1965) suggested that the aspirated stop might also be associated with a greater articulatory tension than the lenis stop, making this factor account for the relatively high pressure of the aspirated stop as well. Finally, Cho & Keating (2001) showed that both aspirated and fortis stops have greater closure duration than lenis stops (as measured by electro-palatography) utterance-initially. This factor might also account for some of the increased pressure in stops in isolation.

3.5.2. *Airflow (Uo)*

We now turn to intraoral airflow (Uo). There was a significant effect of Stop Category ($F(1.49, 14.87) = 49.39, p < 0.0001$) on Uo. But neither a main effect of Dialect nor a Stop Category by Dialect interaction was found. Bonferroni/Dunn pairwise *post hoc* comparisons showed that the fortis stop is produced with substantially less Uo than the lenis stop, consistent with Dart (1987). It also showed that the lenis stop is produced with significantly smaller Uo than the aspirated stop, as shown in Fig. 9(b).

The fact that the fortis stop has greater oral pressure and possibly greater subglottal pressure, but little flow seems counterintuitive, as, all else being equal, the higher the Po, the greater the Uo, as noted by Dart (1987). This seemingly unbalanced high Po and low Uo pattern in fact falls out from a combination of several physiological/aerodynamic factors such as the glottal impedance (which depends largely on the glottal area, Ag), the vocal tract wall tension, and a possible increase in the subglottal pressure (Ps), all of which are crucial aerodynamic parameters in modeling the fortis stop in Dart (1987). As discussed earlier, the greater tension on the vocal tract wall and the heightened Ps associated with the fortis stop would have the effect of heightening Po. However, the stiffened vocal tract wall might also result in a decrease in the volume velocity of airflow since the reduction in the elasticity of the walls would also reduce the amount of elastic recoil of the walls, which would in turn cut down the initial flow volume velocity at release. In addition, a smaller glottal area (and therefore increased glottal resistance) would result in a decrease of transglottal airflow (Ug) and therefore a decrease of Uo. Contrastively, however, the aspirated stop, which has as high Po as the fortis stop, is produced with a far greater Uo, crucially because of a large glottal area at the oral release. This suggests that the effect of the vocal tract wall tension of the aspirated stop assumed by C.-W. Kim (1965) is presumably too small, if existent, to override the effect of the glottal opening. In fact, while greater vocal tract wall tension may, all else being equal, contribute to the reduction of intraoral airflow, this effect would be substantial only if there were only a small transglottal airflow, which is not the case here.
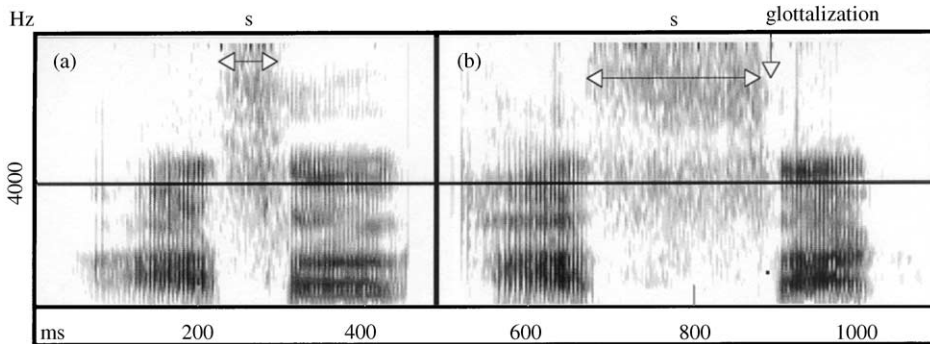
4. Results and discussion for fricatives, /s/ and /s*/

4.1. *Spectrographic observation*

Some acoustic characteristics for /s/ and /s*/ can be qualitatively observed in the spectrograms. Spectrograms of word-initial /s/ and /s*/ are given in Fig. 10. The plane
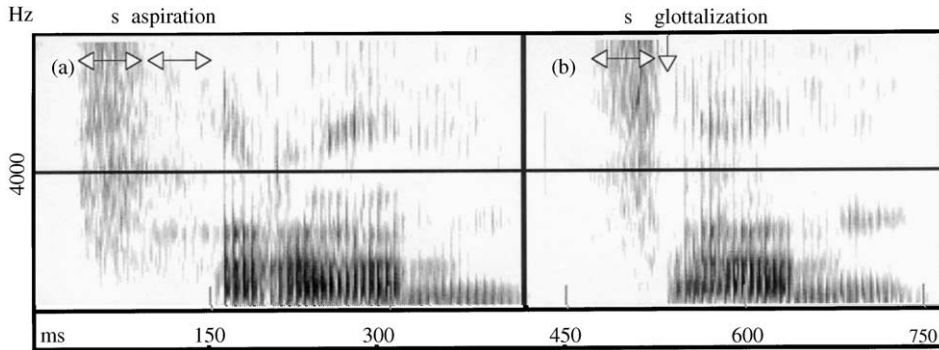
**Figure 10.** Spectrograms of /s/ and /s*/ in word-initial positions (speaker S6): (a) plain /s/, and (b) fortis /s*/.


Figure 11. Spectrograms for /s/ and /s*/ word-medially between vowels: (a) plain /s/ in V_V, and (b) fortis /s*/ in V_V.

/s*/ has two components, frication and aspiration, like a complex, aspirated segment. This complex nature of the plain /s/ was observed in most of the tokens examined across dialects. This suggests that the plain /s/ is phonetically aspirated, agreeing with the previous studies (e.g., Kagaya, 1974; Jun, 1993; Park, 1999). However, it does not always behave like the aspirated stops in that it does not retain a period of aspiration word-medially. The plain /s/ is unaspirated when it occurs word-medially, as shown in Fig. 11(a). In these circumstances, it has intervocalic weakening, like the lenis stops. Furthermore, although it is commonly assumed that /s/ in general does not become voiced intervocalically (which is generally used as evidence that /s/ does not belong to the lenis category, cf. Iverson 1993*a*), we in fact observed that about 46% of tokens of /s/ were fully voiced in this position. Two examples of the fully voiced /s/ between vowels are given in Fig. 12.

A noteworthy point regarding the intervocalic voicing of /s/ is that voicing occurs in a gradient fashion. Fig. 13 shows the percentage of occurrence as a function of the percentage of voicing during the intervocalic /s/. For example, the leftmost bar in the figure indicates that the number of /s/ tokens in which about 15–20% of the fricative duration is voiced (as indicated by the value of the horizontal axis) is about 3% of the
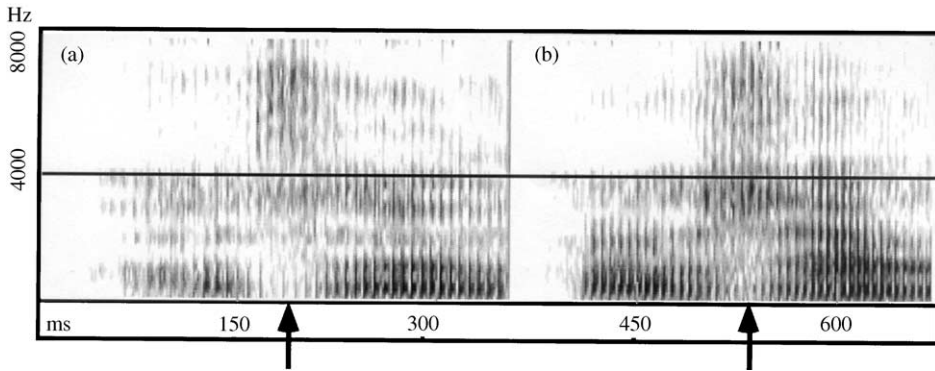
Figure 12. Spectrograms for fully voiced intervocalic /s/: (a) voiced /s/ (spkr S2) and (b) voiced /s/ (spkr S8).
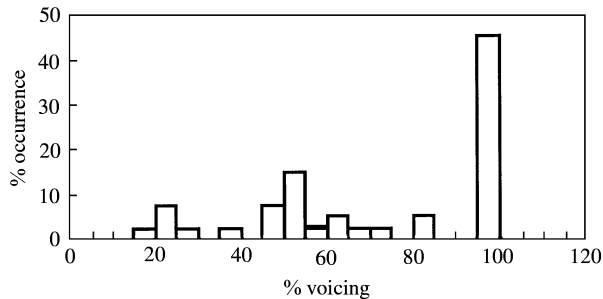

Figure 13. %-voicing during the intervocalic /s/.

entire /s/ tokens. Likewise, the rightmost bar indicates that the number of /s/ tokens with 95–100% voicing throughout the frication is about 47% of the entire /s/ tokens. As can be seen in the figure, there is no clear-cut difference between voiced and voiceless for /s/ in the distribution of the percentage of voicing, but rather, voicing for /s/ occurs over a wide range varying in degree. This gradient nature of intervocalic voicing for /s/ is consistent with Jun's (1995) claim that lenis stop voicing in Korean occurs gradiently (see also J.-I. Han, 2000).

The fortis /s*/ also has two components. As shown in Figs. 10(b) and 11(b), it has a frication portion similar to that of the plain /s/, followed by a vertical gap immediately before the vowel onset. The gap appears to be an indication of the glottalization for /s*/. Fig. 14 shows examples of glottalization extending over the first several cycles of the following vowel. There are a total of 12 out of 24 tokens (50%) for Seoul Korean and nine out of 16 tokens (56%) for Cheju Korean showing a clear gap in the word-initial position. A similar gap is found but less frequently for word-medial tokens in four out of 24 tokens (16%) for Seoul Korean and in eight out of 16 tokens (50%) for Cheju Korean.

Such a glottalization gap was not generally observed when we examined fortis stops. In fortis stops the vocal folds have been shown to be closed just before the oral release and start vibrating immediately for the following vowel, as implied by a short VOT. However, as noted by Jun *et al.* (1998), for the Korean fortis fricative (as well as the plain one) the glottis must remain open to a certain degree in order to permit sufficient airflow
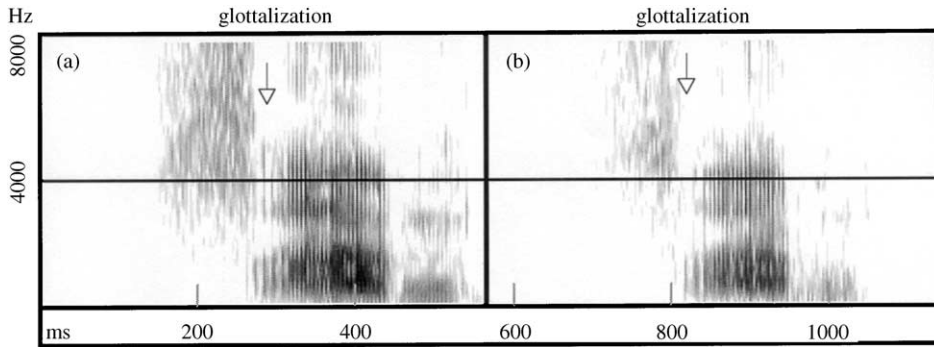
Figure 14. Spectrographic evidence of glottalization for /s*/ in word-initial positions in s*amtʃə, produced twice by speaker C4.

for the frication noise. Consequently, it seems that it is only in the last part of /s*/ that the glottal characteristics of a fortis consonant are present.

4.2. *Fricative duration*

In word-initial position, the mean duration of the plain /s/ (frication plus aspiration) is longer than that of the fortis /s*/ as shown in Fig. 15. This is the opposite of the difference in closure duration between the lenis and the fortis stops; that is, the closure of the fortis stop is longer than that of the lenis stop in word initial position (Cho & Keating, 2001). However, if we exclude the aspiration period associated with the plain /s/, /s*/ becomes significantly longer than /s/ for Seoul Korean ($t(46) = 5.99$, $p < 0.0001$) which is consistent with Yoon's (1998) finding. But there is no such difference in Cheju Korean. A repeated measures ANOVA (Fricative Category by Dialect) showed that the fricative duration is significantly influenced by Fricative Category—i.e., /s*/ is longer than /s/ in the word-initial position ($F(1, 10) = 6.04$, $p < 0.05$). But there is neither a Dialect effect not a Fricative Category by Dialect interaction, showing that the potential dialectal difference may be due simply to chance or speaker variation.

A three-way repeated measures ANOVA (Fricative Category by Position-in-Word (initial/medial) by Dialect) confirmed that while there was a significant effect of Fricative Category ($F(1, 10) = 137.68$, $p < 0.0001$), neither effects of Position-in-Word and Dialect nor an interaction between Dialect and other factors was found. However, there is an interaction between Fricative Category and Position-in-Word ($F(1, 10) = 11.34$, $p < 0.01$). As can be seen in Fig. 15, the Fricative Category by Position-in-Word interaction comes from the fact that the duration of the fortis /s*/ increases from the word-initial to word-medial positions, whereas there is a significant decrease in duration for the plain /s/. In other words, the duration difference between /s/ and /s*/ is even larger in word-medial position. This augmented duration difference in word-medial position can be accounted for by two opposite processes: (1) intervocalic reduction of the plain /s/, which is primarily due to the absence of the aspiration and partial or full voicing and (2) intervocalic enhancement of a geminate nature for the fortis /s*/, which is phonetically realized with a longer duration intervocalically (see Cho & Keating (2001) for the closure duration difference between word-initial and word-medial fortis stops in Korean).
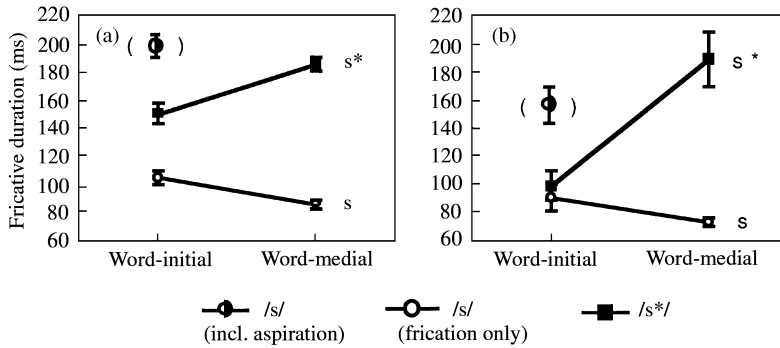
Figure 15. Duration of the plain /s/ and the fortis /s*/ in word-initial and word-medial positions. For the word-initial plain /s/, the duration with and without aspiration are shown. The duration including both frication and aspiration is given in parenthesis. Error bars indicate standard errors. (a) Seoul and (b) Cheju.
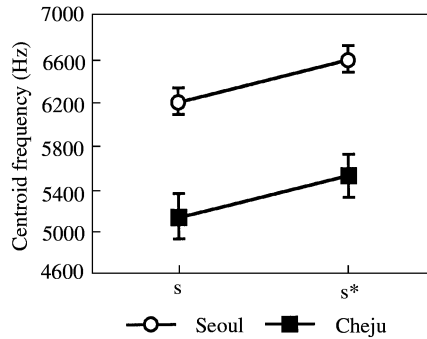
4.3. *Centroid frequency*

There is a small but significant difference in centroid frequency between /s/ and /s*/. A repeated measures ANOVA (Fricative Category by Dialect) showed significant effects of both Fricative Category ($F(1, 10) = 7.25$, $p < 0.05$) and Dialect ($F(1, 10) = 7.99$, $p < 0.05$). However, there was no Fricative Category by Dialect interaction. As can be seen in Fig. 16, /s*/ has a higher centroid frequency than /s/ for both dialects. This may be because /s*/ is produced with a relatively smaller front cavity, which presumably results in a higher centroid frequency. It is also conceivable that the channel between the palate and the tongue necessary for the frication noise of /s*/ may be smaller and thus produce a higher velocity jet of air, which, on striking the upper incisors, would produce a higher frequency. This is consistent with S. Kim (2001) who found a greater linguopalatal contact with a narrower channel for /s*/ than for /s/. Finally, there may be a higher velocity jet of air (and hence a higher frequency) because of a higher subglottal pressure in /s*/ relative to /s/.

The dialect difference comes from the fact that Seoul speakers in general produce /s/ and /s*/ with a higher centroid frequency than Cheju speakers. Cluster analysis showed that seven Cheju speakers and one Seoul speaker were grouped together, indicating that most Cheju speakers behave similarly. One possible explanation for such potential dialectal difference may be that the fricatives /s/ and /s*/ are denti-alveolar for Seoul speakers but alveolar for Cheju. Of course, a cross-dialect articulatory study is needed to verify this possibility.

4.4. *Fundamental frequency (f0)*

A repeated measures ANOVA (Fricative Category by Position-in-Vowel (onset/mid-point) by Dialect) showed that there was a significant effect of Fricative Category ($F(1, 10) = 7.64$, $p < 0.05$) on $f0$, but no effects of Position-in-Vowel and Dialect were found. There was no interaction between factors, except for a non-significant trend for Fricative Category by Dialect interaction ($F(1, 10) = 4.263$, $p = 0.67$). Fig. 17 illustrates effects of Fricative Category and Fricative Category by Dialect interaction. *Post hoc* comparison by Bonferroni/Dunn shows that $f0$ is generally higher for /s*/ than for /s/ at

Figure 16. Centroid frequencies for /s/ and /s*/ in word-initial positions.
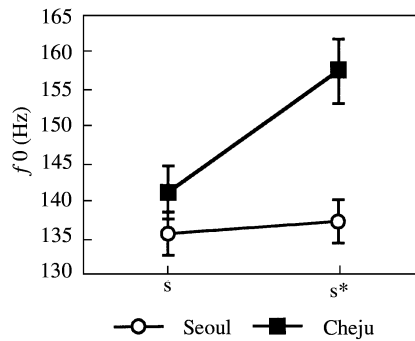
Figure 17. $f0$ difference between /s/ and /s*/ in the following vowel. Error bars indicate standard errors. Data were pooled across the factor Position-in-Vowel (onset *vs.* midpoint) because it has no effect.

$p < 0.1$. The non-significant trend of the Fricative Category by Dialect interaction appears to be due primarily to the fact that the $f0$ of the Cheju /s*/ is significantly higher than that of the Cheju /s/ ($p < 0.01$), whereas there is no such difference for Seoul Korean. However, when the data were observed separately for each speaker, seven Cheju and three Seoul speakers have a higher mean $f0$ for /s*/ than /s/, which suggests that the interaction between Fricative Category and Dialect as a non-significant trend is not primarily due to the dialectal difference but due to speaker variation.

Considering that $f0$ after fortis stops is in general lower than $f0$ after aspirated stops but higher than $f0$ after lenis stops, higher $f0$ after /s*/ than /s/ may suggest that /s/ would belong to a lenis category. When we compared differences in $f0$ between the fricatives and stops, it was found that, in general, $f0$ in the vowel onset after /s/ falls between the lenis and aspirated stops: the mean $f0$ value (137.4 Hz, S.D. = 12.4) after /s/ is higher than after a lenis stop (125.7 Hz, S.D. = 11.6), but lower than after an aspirated stop (147.2 Hz, S.D. = 16.4). On the other hand, $f0$ after the fortis /s*/ (145.3, S.D. = 16.7 Hz) is similar to that after the fortis stop (144.1 Hz, S.D. = 16.3).

4.5. H1−H2 and H1−F2

A repeated measures ANOVA (Fricative Category by Position-in-Vowel by Dialect) indicated that $H1-H2$ is influenced by Fricative Category ($F(1, 10) = 59.48$, $p < 0.0001$),
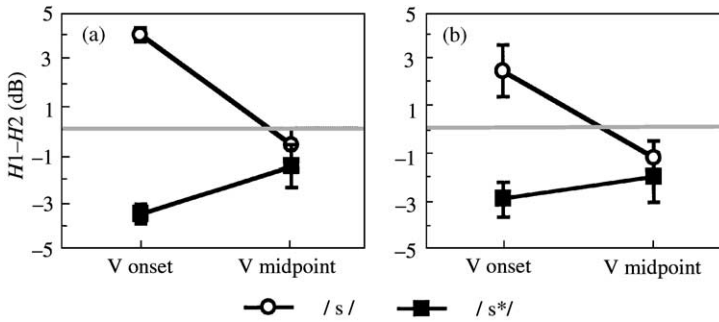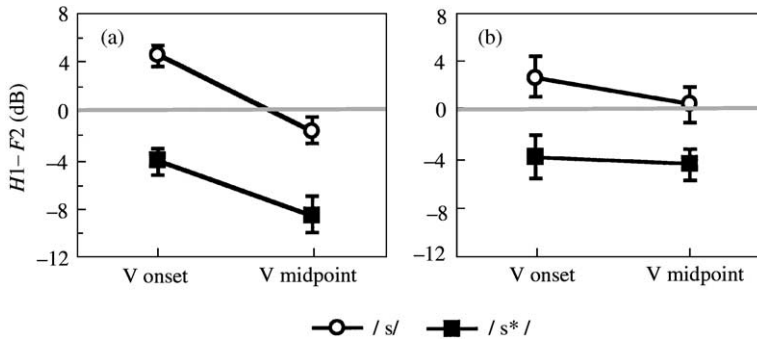
Figure 18. Difference in dB between the first and the second harmonic ($H1-H2$) at the onset and midpoint of the vowel following fricatives. Error bars indicate standard errors. (a) Seoul and (b) Cheju.

but not by Position-in-Vowel and Dialect. There was a significant interaction between Fricative Category and Position-in-Vowel ($F(1, 10) = 17.66$, $p < 0.005$), but no other interactions were found. *Post hoc* comparisons showed that $H1-H2$ is significantly higher for /s/ (positive) than for /s*/ (negative). Fig. 18 illustrates the difference. The difference, however, becomes non-significant at the midpoint of the vowel, which is the major source of the significant interaction between Fricative Category and Position-in-Vowel.

The systematic $H1-H2$ difference between /s/ and /s*/ indicates that at least at the beginning of the vowel, the open quotient is relatively larger for /s/, resulting in breathy phonation, but relatively smaller for /s*/, resulting in pressed phonation. This difference is consistent with that between the lenis and fortis stops as was shown in Fig. 6.

Now let us consider $H1-F2$, from which we can infer the relative abruptness of the vocal folds closing gesture for voicing cycles. A repeated measures ANOVA (Fricative Category by Position-in-Vowel by Dialect) indicated that $H1-F2$ is significantly influenced by Fricative Category ($F(1, 10) = 33.81$, $p < 0.0001$) and Position-in-Vowel ($F(1, 10) = 9.01$, $p < 0.05$) but not by Dialect. There were no significant interactions between factors. A crucial point here is that, in parallel with the $H1-H2$ difference, $H1-F2$ is significantly greater for /s/ than for /s*/, as confirmed by *post hoc* comparisons. Fig. 19 illustrates the difference. As shown in the figure, this difference is maintained through the midpoint of the vowel significantly in both dialects. Put differently, as is the case with vowels after fortis stops, vowels after the fortis /s*/ are generally produced with more abrupt vocal fold closure than vowels after the plain /s/, and such abruptness continues into the middle of the vowel.

It is also interesting to note as to whether $H1-H2$ and $H1-F2$ values for /s/ are closer to those for the aspirated stop or to those for the lenis stop. The mean $H1-H2$ for the plain /s/ (3.4 dB, S.D. = 2.9) is closer to that for the alveolar lenis stop (4.3 dB, S.D. = 3.1) than that for the alveolar aspirated stop (1.4 dB, S.D. = 2.1). But, given the substantial standard deviations, we could infer that there is only a weak tendency towards $H1-H2$ values for /s/ being closer to those for the lenis stop than those for the aspirated stop. As for $H1-F2$, however, even such a tendency could not be observed due to the lack of difference in $H1-F2$ between the alveolar lenis and aspirated stops themselves (cf. Fig. 8).

Figure 19. Difference in dB between the first harmonic and the second formant ($H1 - F2$) at the onset and midpoint of the following vowel. (a) Seoul and (b) Cheju.

5. Summary and general discussion

In the current study, we have observed several acoustic/aerodynamic parameters in accounting for the three-way contrast in stops and the two-way contrast in fricatives. Among many other findings, there are a number of points worth recapitulating. Firstly, stops are systematically differentiated by the voice quality of the following vowel. Secondly, stops are also differentiated by aerodynamic mechanisms. The aspirated and fortis stops are similar in supralaryngeal articulation, but employ a different relation between intraoral pressure and flow. Thirdly, our study suggests that the fricative /s/ has acoustic features that are associated with both lenis and aspirated categories. Finally, although Cheju and Seoul Korean are not mutually intelligible, acoustic and aerodynamic properties of Cheju consonants are very similar in many respects to those of the standard Korean dialect. In what follows, we will first summarize the general findings. Then we will discuss three issues that we believe are worth addressing in conjunction with the data provided by the current study: (1) the role of the following vowel in differentiating the three-way distinction of stops; (2) the phonetic/phonogical nature of /s/; and (3) the implications of the data for theories concerning voicing contrasts and phonological representation of three-way contrastive stops.

5.1. *Summary of results*

In general, most parameters (e.g., VOT, Relative Burst Energy, $f0$, $H1 - H2$ and $H1 - F2$, intraoral air pressure and airflow for stops; centroid frequency, $f0$, $H1 - H2$ and $H1 - F2$ for fricatives) show similar systematic patterns across dialects. In other words, there were neither significant effects of Dialect nor interactions between Dialect and other factors on these parameters, except that VOT tends to be longer in Seoul Korean than in Cheju Korean, especially in the case of the lenis stop. Although Cheju is said to be unintelligible to other Korean speakers, we found that the phonetic properties of obstruents in both dialects are surprisingly similar. We infer that the unintelligibility of Cheju comes in larger part from other components of grammar such as morphosyntactic and/or lexical differences. (See also Cho *et al.* (2001) for dialect differences in the vowel system.) Below is a summary of general patterns primarily based on results of repeated measures ANOVAs and *post hoc* comparisons. In considering the summary, it should be noted that while the

general patterns summarized below are true for a majority of speakers, sometimes, as we observed in the results section, speakers did not show exactly the same pattern for all parameters. Nevertheless, the inferential statistics using repeated measures ANOVAs take into account all speaker deviations from general patterns, so that only truly significant findings are noted in this summary.

5.1.1. *Stops*

a. VOT is longest for the aspirated stop, intermediate for the lenis stop, and shortest for the fortis stop. There is a potential dialectal difference in that there is a non-significant trend toward generally longer VOT in Seoul Korean than in Cheju Korean, which is specifically true for the lenis stop.
b. Burst energy is greater for the aspirated stop than for the lenis and fortis stops, whereas there is no significant difference between the latter two.
c. $f0$ for the lenis stop is always lower than for the fortis and the aspirated stop.
d. $H1-H2$ and $H1-F2$ are both highest for the lenis stop and lowest for the fortis stop, showing that the vowel following the lenis stop is breathier and has a less abrupt adduction phase of the voicing cycle while the vowel following the fortis stop is more pressed and more abrupt in the adduction phase of the voicing cycle.
e. The order of intraoral air pressure (Po) is {aspirated = fortis} > lenis while the order of the intraoral airflow (Uo) is aspirated > lenis > fortis.

5.1.2. *Fricative /s, s*/*

f. The plain /s/ has some aspiration word-initially but no aspiration word-medially.
g. The plain /s/ generally undergoes intervocalic voicing word-medially, in contrast to what has commonly been assumed. About 47% of the tokens have fully voiced /s/ and about another 40% of the tokens are voiced over more than half the frication period.
h. The fortis /s*/ is frequently associated with glottalization (in more than 50% of the tokens) as indicated by a vertical gap in the spectrogram immediately before the vowel onset and irregular voicing at the beginning of the vowel.
i. Centroid frequency of the fortis /s*/ is generally higher than that of /s/, suggesting that the front cavity is relatively smaller in front of the constriction for /s*/ than /s/, and/or that the velocity of the jet of air striking the upper incisors is greater.
j. $f0$ is generally higher for /s*/ than for /s/.
k. $H1-H2$ and $H1-F2$ are both higher for the vowel following /s/ than for the one following /s*/.

### 5.2. *Vowel correlates of the three-way contrast in stops*

Some of the acoustic correlates (e.g., $f0$ difference, $H1-H2$, $H1-F2$) of the three-way contrast of stops examined in the current study are realized on the following vowels. This is not new or surprising in principle. Quite a few studies have suggested that the onset of the following vowels differ in $f0$ (e.g., Han & Weitzman, 1970; Hardcastle, 1973; Jun 1993; Cho, 1996; N. Han, 1998) and intensity buildup (e.g., Han & Weitzman, 1970), which indicates the importance of the vowel in characterizing the three-way contrast in Korean stops. In the current study, however, we found that such differences are in fact not confined to the onset of the following vowel but may spread into the middle of the following vowel.

Jun (1996*b*) claims that the substantial difference in *f*0 between the lenis and other stops cannot be understood simply in terms of the phonetic pitch perturbation caused by the voicing of the preceding consonant (i.e., microprosody) as found in many other languages (Hombert, 1978; Hombert *et al.*, 1979). First, the *f*0 difference between the groups shown in Fig. 5 is substantially greater than that found in other languages (e.g., Kohler, 1982; Kingston & Diehl, 1994), and the *f*0 difference is still present at the midpoint of the syllable. Furthermore, all three stops are voiceless unlike other languages where *f*0 differences are associated with the voiced *vs.* voiceless distinctions. Jun claims that this segmentally triggered *f*0 distinction of Seoul Korean is phonologized in the intonation system of Seoul Korean as described by Jun (1993, 1996*a*, 1998). In this intonation system, the Accentual Phrase is realized on two distinct tonal patterns, LHLH and HHLH, where the phrase-initial tone (L *vs.* H) is determined by the type of the consonant occurring in the onset of the phrase initial syllable. When there is a fortis or aspirated obstruent in the onset, the Accentual Phrase starts with H, and otherwise (i.e., when there is a lenis obstruent, a sonorant consonant, or a vowel), it starts with L. This phonologization of microprosody is confirmed by our data from Seoul, and seems to occur in Cheju as well, though we do not yet know the details of the intonation system of Cheju.

Considering $H1 - H2$ and $H1 - F2$ to be indicative of phonation types, one of the findings in the current study is that the voice quality of the following vowel changes systematically as a function of preceding stop categories, and that this difference is maintained in the middle of the vowel, though it becomes smaller. Languages like Mazatec and Mpi where vowels have either two-way or three-way phonological contrasts in phonation types (i.e., modal, breathy, and pressed) do not in fact show acoustic differences for the entire time course of the vowel (Blankenship, 1997, 2002). For example, Silverman, Blankenship, Kirk & Ladefoged (1995) reported that in Mazatec, a difference in breathiness is maintained only up to 43% of the vowel, leaving modal phonation for the rest of the vowel. Korean is a case where phonation differences in the vowels do not occur contrastively as they do in Mazatec, but rather phonetically due to the consonant context. Nonetheless, our data show quite robust differences even in the middle of the vowel, suggesting that the voice quality of the vowel plays an important role in differentiating the three-way contrast.

The importance of vowel correlates in the three-way contrast in stops is illuminated by a perception experiment reported in Cho (1995, 1996). In this experiment, Korean speakers heard only the vowels extracted from recorded CV syllables in which C was one of the Korean stops. Subjects were instructed to identify the type of the missing preceding stop by choosing one of three categories (lenis, fortis, and aspirated) knowing only the place of articulation and the vowel. For example, one set of the choices looked like (1) *pa* (2) *pʰa* (3) *p*a* (in the original sheet the choices were typed in Korean orthography), and subjects chose one of these categories by listening to the extracted vowel portion. Cho found that although different listeners responded to the stimuli somewhat differently, the overall accuracy rate was 67%. (The chance level was 33% with a forced choice out of three categories.) Interestingly, as shown in Fig. 20, listeners perceived lenis stops and fortis stops with a greater accuracy rate (71–83% and 65–86%, respectively) than aspirated stops (42–57%), from which one could infer that listeners without access to VOT differences are cued by the nonmodal phonation types (i.e., the breathy and the pressed phonations for the lenis and fortis stops, respectively). Furthermore, *post hoc* comparisons among correctly perceived and misperceived tokens showed that aspirated
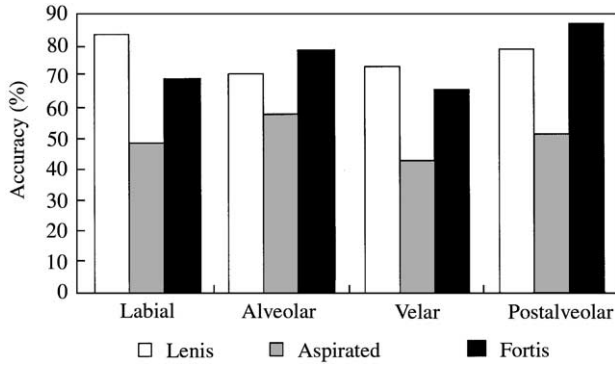
Figure 20. Accuracy rates in identifying the type of a missing preceding stop from the sound of the following vowel extracted from CV syllables (redrawn from Cho, 1996, p. 73). The accuracy rates are averaged across ten subjects who participated in the perception test.

consonants were more often misperceived as fortis consonants (52%) than as lenis ones (10%), and fortis consonants were more often perceived as aspirated consonants (26%) than as lenis consonants (12%). Jun (1996*b*) interprets this pattern of error as a reflection of $f0$ differences in that both fortis and aspirated stops have a higher $f0$ in the following vowel, and therefore are interchangeably misperceived as each other.

Although further systematic perception tests are required to know exactly which acoustic/perceptual cues embedded in vowels are utilized by listeners, it seems safe at the moment to conclude that the acoustic and perceptual information of the preceding consonants spreads to the following vowels to a large extent, and thus helps listeners identify preceding stop categories (see Kim, Beddor & Horrocks (2002) for a recent study about the contribution of vocalic information to the perception of Korean initial stops). This may explain the results of Abramson and Lisker's (1973) perception study where only VOTs were manipulated. They found that with VOTs alone, speakers do not consistently differentiate among the three-way contrastive stops in Korean. We expect that stops would be better differentiated when speakers are presented with both VOT and vowel quality information. In future tests, it will be interesting to see the extent to which the Korean stops can be differentiated by listeners presented with various acoustic parameters while keeping others intact.

5.3. *Phonetic/phonological nature of plain /s/*

We found some phonetic properties that may contribute in distinguishing the contrasting segments /s/ and /s*/ in Korean. The phonetic nature of /s*/ as fortis is rather straightforward: it is produced with (1) frication and frequent glottalization immediately before the vowel, (2) relatively higher centroid frequency, (3) higher $f0$ in the vowel onset, (4) pressed voicing (negative $H1 - H2$) at the vowel onset, and (5) abrupt vocal fold closure during the following vowel (negative $H1 - F2$).

The phonetic and phonological nature of the plain /s/ is quite complex and ambiguous. It is produced with frication, aspiration, and $f0$ at the vowel onset that are relatively high as compared to a lenis stop. This might make one believe that it is an aspirated /s/ appropriately transcribed as [s^h]. However, we also found several reasons as to why the

plain /s/ might be categorized as a lenis segment. First, its counterpart lenis stop is also produced with a fair amount of aspiration. Second, the onset of the vowel after /s/ has a similar breathy voice quality to vowels after the lenis stop. Third, $f0$ in the vowel onset after /s/, though higher compared to $f0$ after a lenis stop, tends to be lower than that after an aspirated stop. Fourth, the plain /s/ loses its aspiration word-medially as does the lenis stop. Finally, though it has been generally assumed that /s/ does not become voiced (cf. Iverson, 1983*a*), we have observed that about 48% of the tokens were fully voiced intervocalically, which undermines the contention that /s/ behaves like an aspirated stop, which always fails to undergo intervocalic voicing. Considering all these observations, it seems that /s/ has more in common with the lenis category than the aspirated one in terms of its phonetic nature. This is in harmony with its phonological behavior as a lenis category: i.e., it forms a natural class with a lenis stop in the Post Obstruent Tensing rule in Korean.

5.4. *Phonological representation and phonetic realization*

What are the implications of our findings for theories of voicing contrast and phonological representation? It is apparent that Korean and Cheju have phonological contrasts that are unlike those in any other language. It has long been known that the Korean fortis stops are unique. Now it appears that the aspirated stops may also be atypical—there are no reports that we know of indicating that any other language has aspirated stops that raise the fundamental frequency of the following vowel to such an extent. In addition, the lenis stops are also unusual in that they are aspirated but become voiced intervocalically, and have a slightly breathy mode of vibration of the vocal folds.

It would take a great deal of procrustean effort to force Korean stops into the categories that have been developed for phonological descriptions of other languages. Many of the languages with three way contrasts have been discussed in terms of the categories [voiced], [voiceless unaspirated] and [aspirated] (cf. Jones, 1950; Lisker & Abramson, 1964, *inter alia*). Keating (1984) discusses the phonetic realization of these categories, proposing a polarization theory, which postulates that the distribution of VOT values (as a measure of voicing) will be maximally distributed within the limits of the particular phonetic categories chosen by the language. In this view languages choose one, or two, or all three of these phonetic categories for their contrasting stops. This makes the theory applicable not only to a language with a three-way contrast in stop voicing such as Thai, but also to a language with no contrast in voicing, such as Hawaiian, or to a two-way contrast in stop voicing such as English. In English, /b d g/ are phonetically realized as voiced in word-medial position, but as voiceless unaspirated in word-initial position, whereas /p t k/ are phonetically realized as voiceless aspirated in stressed syllables but as voiceless unaspirated either in unstressed syllables or when they are preceded by a fricative as in *star*.

However, as Keating has noted, there is no way that such a theory could be applied to Korean, which has a three-way contrast in stops, all of them being voiceless in word- or phrase-initial position. We would have to add another possibility to Keating's three categories, leading to [voiced], [voiceless unaspirated], [voiceless slightly aspirated], and [voiceless heavily aspirated]. Now the question is whether we need all these four phonetic categories or not. If we consider only VOT differences, then four phonetic categories seem inevitable. This undermines Keating's proposal that no language needs more than three voicing categories. However, it may be the case that Korean is

a language in which these phonetic categories may not be the appropriate phonetic representations that characterize stops. For example, consider all the phonetic correlates of these stops that are associated with vowels. Vowels are breathy after the lenis stop or pressed after the fortis stop; their $f0$ is perturbed to a greater degree than is found for phonetic perturbation in other languages; and they have different characteristics of intensity buildup. Auditory information from just the vowels can help listeners identify preceding consonant types at a far greater than chance level. The various phonetic correlates appear to suggest that phonetic voicing categories that characterize stops in terms of VOT may not be the appropriate phonetic representations for Korean.

Then, the question is how Korean stops can be specified by phonological features. One generally assumed featural specification of Korean stops is found in Halle & Stevens (1971), where combinations of two laryngeal features [±Spread Glottis] and [±Constricted Glottis] reflect phonetic dimensions of the glottal width that are varied with stop categories, as can be seen in the first two rows of Table VI. However, this does not reflect another important phonetic dimension, the amount of the tension in the vocal folds (cf. C.-W. Kim, 1965; Hirose *et al.*, 1974). Halle & Stevens (1971) also suggested two other laryngeal features, [±Stiff Vocal Folds] and [±Slack Vocal Folds] to characterize the state of larynx at the moment of release. Iverson (1983b) suggested using [±Stiff Vocal Folds] in order to take the vocal fold tension into account, as can be seen in Table VI. Adding the feature [±Stiff Vocal Folds] to the system, seemingly redundant from the perspective of traditional phonology, differs from the Halle and Stevens' system in that it allows for a closer link between phonetics and phonology. The stiffness information allows us to predict various phonetic correlates associated with stops and the following vowel. For example, higher $f0$ and the abrupt intensity buildup at the onset of the vowel is predicatable for the fortis stop from [+Stiff Vocal Folds] associated with it. The stiffness information also captures the grouping of aspirated and tense stops against lenis stops in terms of their behavior with respect to $f0$ and intervocalic voicing. Thus, it seems that the feature system in Table VI illuminates a closer phonetics/phonology relationship. However, we cannot use it to describe the phonetic nature of the lenis stop voicing in phrase- or word-medial position because a binary feature system assumes a categorical phonological rule and the presence/absence of a feature categorically.

Another feature system which works better from the perspective of phoneticians is a system in which features are privative in the framework of underspecification theory. Lombardi (1991, 1995) proposed that the three-way contrastive stops in Korean could be phonologically differentiated by two privative laryngeal features, Spread Glottis and

TABLE VI. Laryngeal features of Korean consonants, based on Iverson (1983*b*) and Halle & Stevens (1971)

Phonological features	Aspirated	Lenis	Fortis
Spread Glottis	+	−	−
Constricted Glottis	−	−	+
Stiff Vocal Folds	+	−	+
Phonetic correlates			
Glottal Width	Large	Intermediate	Small
Glottal Tension	High	Low	High

Note: The lower part of the table shows possible phonetic dimensions.

Constricted Glottis. That is, as given in (1), the aspirated and fortis stops are specified as Spread Glottis and Constricted Glottis, respectively, reflecting the phonetic dimension of the glottal width: Spread Glottis refers to a wide glottal opening and Constricted Glottis to a narrow one. The lenis stop is unspecified for any of these features, and is phonetically realized with a neutral vocal fold configuration, resulting in an intermediate size of the glottal opening. We further assume that the phonetic dimension of the vocal folds tension can be stated in phonology by redundancy rules (cf. Steriade, 1987) as stated in (2), whereby a lenis stop, which is not specified for glottal state, does not get stiff vocal folds.

(1) fortis stop   [constricted glottis]
    aspirated stop   [spread glottis]
    lenis stop   unspecified

(2) redundancy rules
    [constricted glottis] → [stiff vocal folds]
    [spread glottis] → [stiff vocal folds]

There are several reasons as to why the privative feature system is preferable to the binary system. Firstly, in the Halle and Stevens' binary system, both the lenis and the fortis stops are specified with [− Spread Glottis]. This does not reflect the difference between the lenis and the fortis stops in terms of glottal width or amount of aspiration. In a privative system with an underspecification framework as stated in (1) and (2), the lenis stop, which is unspecified for laryngeal features, is associated with a neutral position of the vocal folds, resulting in an intermediate glottal width. Secondly, redundancy rules as stated in (2) may be viewed as language-specific. Stiff Vocal Folds may be phonetically realized as having higher $f0$ at the vowel onset after the aspirated and fortis stops. This is similar in spirit to Stevens *et al.*'s (1986) featural enhancement in that secondary or redundant features reinforce phonetic contrasts. Finally, another advantage in having the lenis stop unspecified for either Spread Glottis or Constricted Glottis is that the underspecification seems to better capture characteristics of the Lenis Stop Voicing rule. Pierrehumbert (1980) and Keating (1988) assume that underspecification may persist in phonetic representations and even at the surface level. One of the clearest examples is the case of /h/, which has no intrinsic oral feature values, and whose formant frequencies are determined by the context, as a result of interpolation between the surrounding vowels. Likewise, we posit that the unspecification of laryngeal features for the lenis stop persists in the phonetic representation. The important point is that the phonetic realization of the lenis stop is context-sensitive, so that when it is surrounded by voiced segments within an Accentual Phrase, it becomes voiced passively, being influenced by the voicing from the surrounding vocalic gesture. Furthermore, experimental evidence shows that the lenis stop becomes voiced in a gradient fashion. The patterns of voicing during the lenis stop closure vary with the level of the prosodic boundary that precedes it (Jun, 1995; Cho & Keating, 2001). Roughly, the amount of voicing decreases as the level of the prosodic boundary increases from a word-boundary to an Accentual Phrase boundary and on to an Intonational Phrase boundary. Jun (1995) attributes this phenomenon to overlapping of the glottal gestures. The prosodic position affects the gestural amplitude and degree of overlapping, resulting in a gradient range of voicing. This leads to her conclusion that the Lenis Stop Voicing Rule is a gradient phonetic rule. Thus, Korean offers a nice example of the use of underspecification. The underspecified feature of the lenis stop reflects

a phonetic dimension (neutral vocal folds position) that persists in the phonetic realization, allowing context-sensitive and passive lenis stop voicing.

From the phonetic evidence that we have, it is obvious that there is no one-to-one mapping between phonological features and phonetic values, but rather various phonetic dimensions are marshalled to implement the phonological contrasts in the language. This clearly necessitates a sophisticated set of phonetic rules in the grammar that implements phonological features so that they are physically realized in a language-specific manner. The underspecification and redundancy rules discussed in previous paragraphs certainly bring phonetics and phonology into a closer relationship. However, as noted by Docherty (1992), all the currently available speech production models such as Keating's (1984) phonetic representations (which bridges the phonological representation and phonetic realization) and Browman & Goldstein's (1986, 1990, 1992) articulatory phonology (which incorporates an abstract temporal dimension in phonological representations) are still not specific enough to characterize language-specific phonetic realizations. Much more work on speech production modeling needs to be done to capture low-level phonetic realizations that are language-specific and not governed by universal phonetic implementation rules.

5.5. *Concluding comment*

In the current study, we have observed several acoustic/aerodynamic parameters in accounting for the three-way contrast in Korean stops and the two-way contrast in Korean fricatives. Some of the results (e.g., VOT, $f0$ at the onset of the vowel, Po and Uo) replicate previous findings, while others (e.g., VOT as a function of place of articulation, stop burst energy, $f0$ and voice quality of the following vowel, Po and Uo for the aspirated stops, centroid frequency for fricatives) contribute new data to the literature about Korean obstruents. The Korean consonant system is unique among the world's languages. It is worthy of intense study in any investigation of the limits of human speech capabilities. The complexity of the phonetic properties of the three-way contrastive Korean stops implies that speech production is far more complicated than we can predict from simplified phonological representations or even the more sophisticated general speech production models that are currently available. We hope that the present study will serve as a substantial reference for future research bearing not only on issues about Korean obstruents and related issues addressing phonation types and voicing contrasts in other languages, but also on fine-grained modeling of speech production in general.

References

Ahn, H. (1999) *Post-release phonatory processes in English and Korean: acoustic correlates and implications for Korean phonology*. PhD dissertation, University of Texas, Austin.

Abberton, E. (1972) Some laryngographic data for Korean stops, *Journal of International Phonetic Association*, **2,** 67–78.

Abramson, A. S. & Lisker, L. (1973) Voice timing in Korean stops. *In Proceedings of the seventh international congress of "phonetic sciences"*, Montreal, pp. 439–446.

Bickley, C. (1982) Acoustic analysis and perception of breathy vowels. *Speech Communication Group Working Papers I*. Research Laboratory of Electronics, MIT, pp. 73–83.

Blankenship, B. (1997) *The time course of breathiness and laryngealization in vowels.* PhD dissertation, UCLA.

Blankenship, B. (2002) The timing of nonmodal phonation in vowels, *Journal of Phonetics*, **30**(2), 193–228.

Browman, C. & Goldstein, L. (1986) Toward an articulatory phonology, *Phonology Yearbook*, **3**, 219–252.

Browman, C. & Goldstein, L. (1990) Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology I*: *Between the grammar and the physics of speech* (J. Kingston & M. Beckman, editors), pp. 341–376. Cambridge, U.K.: Cambridge University Press.

Browman, C. & Goldstein, L. (1992) Articulatory phonology: an overview, *Phonetica*, **49**, 155–180.

Childers, D. G. & Lee, C. K. (1991) Vowel quality factors: analysis, syntheses, and perception, *Journal of the Acoustical Society of America*, **90**(5), 2394–2410.

Cho, T. (1995) Korean stops and affricates: acoustic and perceptual characteristics of the following vowel, *Journal of the Acoustical Society of America*, **98**(5), 2891 [Abstract].

Cho, T. (1996) *Vowel correlates to consonant phonation*: *an acoustic-perceptual study of Korean obstruents.* MA thesis, University of Texas at Arlington.

Cho, T. (1998) Domain-initial strengthening in the prosodic hierarchy: an EPG study. In *Proceedings of 11th international conference on "Korean linguistics (ICKL 11)"*, July 6–9, University of Hawaii at Manoa.

Cho, T., Jun, S.-A., Jung, S. & Ladefoged, P. (2001) The vowels of Cheju, *Korean Journal of Linguistics* (Ean-Eo), **26**(4), 801–816. (An earlier version appeared in *UCLA Working Papers in Phonetics*, **98**, 81–94.)

Cho, T., Jun, S.-A., Ladefoged, P. (2000) An acoustic and aerodynamic study of consonants in Cheju, *Speech Sciences*, **7**, 109–142. (*The Journal of the Korean Association of Speech Sciences.*) (Also in *UCLA Working Papers in Phonetics*, **98**, 54–80).

Cho, T. & Keating, P. (2001) Articulatory and acoustic studies of domain-initial strengthening in Korean, *Journal of Phonetics*, **29**(2), 155–190. (An earlier version appeared in *UCLA Working Papers in Phonetics*, **97**, 100–138.)

Cho, T. & Ladefoged, P. (1999) Variations and universals in VOT: evidence from 18 languages, *Journal of Phonetics*, **27**(2), 207–229.

Cho, Y. Y. (1990) Syntax and Phrasing in Korean. In *The phonology–syntax connection* (S. Inkelas & D. Zec, editors), pp. 47–62. Chicago: University of Chicago Press.

Dart, S. (1987). An aerodynamic study of Korean stop consonants: measurements and modeling, *Journal of the Acoustical Society of America*, **81**(1), 138–147.

Docherty, G. (1992) *The timing of voicing in British English obstruents.* Berlin: Foris.

Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R. (1988) Statistical analysis of word-initial voiceless obstruents: preliminary data, *Journal of the Acoustical Society of America*, **84**(1), 115–123.

Halle, M. & Stevens, K. N. (1971) A note on laryngeal features, *Quarterly Progress Report* (Research Laboratory of Electronics, MIT), **101**, 198–212.

Han, J.-I. (2000) Intervocalic stop voicing revisited, *Speech Sciences*, **7**, 203–216. (*The Journal of the Korean Association of Speech Sciences.*)

Han, J.-I. (1996) *The phonetics and phonology of "tense" and "plain" consonants in Korean.* PhD dissertation, Cornell University.

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

Han, N. (1998) *A comparative acoustic study of Korean by native Korean children and Korean–American children.* MA thesis, UCLA.

Hardcastle, W.J. (1973) Some observations of the tense-lax distinction in initial stops in Korean, *Journal of Phonetics*, **1**, 263–271.

Hays, W. L. 1994. *Statistics*, 5th edition. New York: Harcourt Brace College Publishers.

Hirose, H., Lee, C. Y. & Ushijima, T. (1974) Laryngeal control in Korean stop production, *Journal of Phonetics*, **2**, 145–152.

Hombert, J.-M. (1978) Consonant types, vowel quality, and tone. In *Tone*: *a linguistic survey* (V. Fromkin, editor), pp. 77–111. New York: Academic Press.

Hombert, J.-M., Ohala, J. & Ewan, W. (1979) Phonetic explanations for the development of tones, *Language*, **55**, 37–58.

Huffman, M. (1987) Measures of phonation type in Hmong, *Journal of the Acoustical Society of America*, **81**, 495–504.

Huynh, H. & Feldt, L. S. (1970) Conditions under which mean square ratios in repeated measurements designs have exact F-distributions, *Journal of the American Statistical Association*, **65**, 1582–1589.

Iverson, G. (1983a) Korean s, *Journal of Phonetics*, **11**, 191–200.

Iverson, G. (1983b) On glottal width features, *Lingua*, **60**, 331–339.

Jones, D. (1950). *The phoneme*: *its nature and use*. Cambridge: W. Heffer & Sons.

Jun, S.-A. (1993) *The phonetics and phonology of Korean prosody*. PhD dissertation, Ohio State University.

Jun, S.-A. (1995) Asymmetrical prosodic effects on the laryngeal gesture in Korean. *Phonology and phonetic evidence*: *papers in laboratory phonology IV* (B. Connell & A. Arvaniti, editors), pp. 235–253. Cambridge, U.K.: Cambridge University Press.

Jun, S.-A. (1996*a*) *The phonetics and phonology of Korean prosody*: *intonational phonology and prosodic structure*. New York: Garland Press.

Jun, S.-A. (1996*b*) Influence of microprosody on macroprosody: a case of phrase initial strengthening, *UCLA Working Papers in Phonetics*, **92,** 97–116.

Jun, S.-A. (1998) The accentual phrase in the Korean prosodic hierarchy, *Phonology*, **15**(2), 189–226.

Jun, S.-A., Beckman M. & Lee, H. (1998) Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents, *UCLA Working Papers in Phonetics*, **96,** 43–68.

Kagaya, R. (1974) A fiberscopic and acoustic study of the Korean stops, affricates and fricatives, *Journal of Phonetics*, **2,** 161–180.

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

Keating, P. A. (1988) Underspecification in phonetics, *Phonology*, **5,** 275–292.

Kim, C.-W. (1965) On the autonomy of the tensity feature in stop classification, *Word*, **21,** 339–359.

Kim, C.-W. (1970) A theory of aspiration, *Phonetica*, **21,** 107–116.

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

Kim, M.-R., Beddor, P. S. & Horrocks, J. (2002) The contribution of consonantal and vocalic information to the perception of Korean initial stops, *Journal of Phonetics*, **30**(1), 77–100.

Kirk, R. E. (1995) *Experimental Design*: *Procedures for the Behavioral Sciences*. Pacific Grove, CA: Brooks/Cole Publishing Company.

Kim, S. (2001) *The interaction between prosodic domain and segmental properties*: *domain initial strengthening of fricatives and post obstruent tensing in Korean*, MA thesis, University of California, Los Angeles.

Kim, Y. (1995) *Acoustic characteristics of Korean coronal*. MA thesis, University of Texas at Arlington.

Kingston, J. & Diehl, R. (1994) Phonetic knowledge, *Language*, **70**(3), 419–454.

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

Klatt, D. & Klatt, L. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers, *Journal of the Acoustical Society of America*, **87,** 820–857.

Kohler, K. J. (1982) F0 in the production of lenis and fortis plosives, *Phonetica*, **39**(4–5), 199–218.

Ladefoged, P. (1964) *A phonetic study of west African languages*. Cambridge: Cambridge University Press.

Ladefoged, P. (1983) The linguistic use of different phonation types. In *Vocal fold physiology*: *contemporary research and clinical issues* (D. Bless & J. Abbs, editors), pp. 351–360. San Diego: College Hill Press.

Ladefoged, P. (1997) Instrumental techniques for linguistic phonetic fieldwork. In *The handbook of phonetic sciences* (W. Hardcastle & J. Laver, editors), pp. 137–166. Oxford: Blackwell Publishers.

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

Ladefoged P., Middieson, I. & Jackson, M. (1987) Investigating phonation types in different languages. In *Vocal physiology*: *voice production, mechanisms and functions* (O. Fujimura, editor), pp. 297–317. New York: Raven.

Lehiste, I. & Peterson, G. E. (1961) Some basic considerations in the analysis of intonation, *Journal of the Acoustical Society of America*, **33,** 419–425.

Lieberman, P. (1963) Some acoustic measures of the fundamental periodicity of normal and pathologic larynges, *Journal of the Acoustical Society of America*, **35,** 344–353.

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

Lombardi, L. (1991) *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts, Amherst.

Lombardi, L. (1995) Laryngeal features and privativity, *The Linguistic Review*, **12,** 35–59.

Max, L. & Onghena, P. (1999) Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research, *Journal of Speech, Language, and Hearing Research*, **42,** 261–270.

Park, H. S. (1999) The phonetic nature of the phonological contrast between the lenis and fortis fricatives in Korean. In *Proceedings of the 14th International congress of* "*Phonetic science (ICPhS* 99)", San Francisco, Vol. 1, pp. 424–427.

Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation. MIT.

Pierrehumbert, J. & Talkin, D. (1992) Lenition of /h/ and glottal stop. In *Papers in laboratory phonology II*: *gesture, segment, prosody* (G. Docherty & D. R. Ladd, editors), pp. 90–117. Cambridge, UK: Cambridge University Press.

Silva, D. J. (1992) *The Phonetics and phonology of stop lenition in Korean*. PhD dissertation, Cornell University.

Silverman, D., Blankenship, B., Kirk, P. & Ladefoged, P. (1995) Phonetic structures in Jalapa Mazatec, *Anthropological Linguistics*, **37**(1), 70–88.

Silverman, D. & Jun, J. (1994) Aerodynamic evidence for articulatory overlap in Korean, *Phonetica*, **51,** 210–220.

Steriade, D. (1987) Redundant values. In A. Bosch, B. Need & E. Schiller (eds.) *Paters from the parasession on metrical and autosegmental phonology*, Chicago Linguistic Society, 23, 339–362.

Stevens, K. N. (1999) *Acoustic phonetics*. Cambridge: MIT Press.

Stevens, K. N., Keyser, S. J. & Kawasaki, H. (1986) Toward a phonetic and phonological theory of redundant features. In *Invariance and variability in speech processes* (J. S. Perkell & D. H. Klatt, editors), pp. 426–463. New Jersey: Lawrence Erlbaum Associates.

Yoon, K. (1998) *An acoustic analysis of Korean aspirated and tense alveolar fricatives using Multi-Speech*. MA thesis, University of Kansas.

Zsiga, E. (1993) *Features, gestures, and the temporal aspects of phonological organization*. PhD dissertation, Yale University.