The status of voicing and aspiration as cues to Korean laryngeal contrast*

Charles B. Chang
University of California, Berkeley

1 Introduction

Previous research on the three-way Korean laryngeal contrast among lenis, fortis, and aspirated plosives has described these series in great detail. These studies have shown that word-initial realizations of the three laryngeal series show a number of phonetic differences with respect to each other. There are differences in linguopalatal contact (Cho and Keating 2001), glottal configuration (C. W. Kim 1970, Kagaya 1974), subglottal and intraoral pressure (Dart 1987), laryngeal and supralaryngeal articulatory tension (C.-W. Kim 1965, Hardcastle 1973, Hirose *et al*. 1974, Dart 1987), voice onset time (Lisker and Abramson 1964, Han and Weitzman 1970, Hardcastle 1973, Hirose *et al*. 1974, J. I. Han 1996, Cho *et al*. 2002, Choi 2002, M. Kim 2004), fundamental frequency onset (J.-I. Han 1996, Choi 2002, M. Kim 2004), first and second formant onset (Park 2002), intensity of vowel onset (Han and Weitzman 1970), and voice quality of vowel onset (N. Han 1998, Cho *et al*. 2002, Kim and Duanmu 2004)[1]. What most of these studies have in common is that they analyze the contrast as being among three series of phonologically voiceless plosives – a contrast that is extremely rare, if not unique, typologically. Although the lenis series typically becomes voiced intervocalically, it, along with the other two series, has traditionally been transcribed with symbols for voiceless sounds: /p, t, k/ for lenis, /p*, t*, k*/ or /p', t', k'/ for fortis, and /p^h, t^h, k^h/ for aspirated. Even when transcription conventions make use of symbols for voiced sounds to represent Korean lenis obstruents, the analysis of the laryngeal contrast usually leaves voicing out of underlying representations. For instance, the transcription conventions of Lee (1999), which are used in many other phonetic studies of Korean, uses /b, d, g/ to indicate "voiceless unaspirated (or slightly aspirated) lenis plosives."

* This work was conducted with the support of a Jacob K. Javits Fellowship from the U.S. Department of Education. I am grateful to Erin Haynes, Larry Hyman, Sharon Inkelas, Keith Johnson, Ian Maddieson, John Ohala, the Berkeley Phonetics and Phonology Forum, and audiences at UW-Madison, UNC-Chapel Hill, and CLS 43 for helpful discussions and feedback. I would also like to thank all the speakers and listeners who were gracious enough to participate in this study. Any remaining errors are mine.

[1] The somewhat complementary dimensions of closure duration and vowel length also appear to be important cues to the laryngeal distinction (cf. Silva 1992, Kim 1994, Han 1996, Cho and Keating 2001), but primarily in postvocalic environments. As this paper focuses on the laryngeal contrast in prevocalic position, these factors are not considered further here, though closure duration does appear to play a role in the perception of initial voiced plosives (cf. §2.2).

On the other hand, Kim and Duanmu (2004) have argued that the Korean laryngeal contrast is not unique and, instead, that it is more accurately characterized as one among voiced stops (which would then be analyzed as devoicing word-initially), voiceless unaspirated stops, and voiceless aspirated stops – a system of laryngeal contrast not uncommon in other languages of the world (e.g. Burmese, Thai, Armenian, Hindi, etc.).

The present study contributes to the growing body of research on this unusual laryngeal contrast with a cross-linguistic perceptual survey of plosives from three other languages with different systems of laryngeal contrast. How do Korean speakers perceive phonation in these other languages in terms of the laryngeal categories of Korean? Do Korean speakers' perceptual patterns provide evidence in favor of a particular analysis of the Korean laryngeal system? What are the acoustic correlates of these perceptual patterns?

This paper is organized as follows. Section 2 describes the design and results of a cross-linguistic perception experiment conducted with native Korean speaker subjects and stimuli from Sindhi, Spanish, and Shanghainese. Section 3 provides acoustic data on these stimuli and statistical analyses cross-correlating the perceptual data with the acoustic data. Section 4 discusses the implications of these results, and Section 5 summarizes the main conclusions.

2 Experiment 1: cross-linguistic perception

2.1 Methods

2.1.1 Materials

In order to collect the stimuli that listeners would provide judgments on, word lists were constructed for Korean, Sindhi, Spanish, and Shanghainese that included the laryngeal contrasts of interest: in Korean and Shanghainese, a three-way contrast between lenis/voiced, fortis/voiceless unaspirated, and voiceless aspirated; in Spanish, a two-way contrast between voiced and voiceless; and in Sindhi, a four-way contrast between voiced unaspirated, voiced aspirated, voiceless unaspirated, and voiceless aspirated (as well as some implosives). Whenever possible, words were included that had the relevant consonants in both word-initial and word-medial positions and that were matched for vowel quality (see Chang 2007 for lists of words used, which are not included here for reasons of space).

2.1.2 Stimuli

Native speakers were recorded saying their respective word lists in a sound-attenuated booth at the University of California, Berkeley, using a Marantz PMD670 solid state recorder and an AKG C420 condenser microphone. Three tokens of each word were recorded in isolation as mono sound files at a sampling rate of 44.1 kHz and a bit rate of 16 bps. Once the words were recorded, individual syllables were extracted from the multisyllabic words in Praat 4.2.17 (Boersma and Weenink 2004). The duration of these excised syllables as well as of the monosyllabic words was normalized to 0.5 seconds. A perceptual test was then constructed mixing 120 originally word-initial and originally non-word-initial syllables from all four languages.

2.1.3 Subjects

The five speakers who recorded the stimuli were all students in their 20s or 30s with no articulatory or auditory impairments. Speaker 1 (L1 = Korean) was a male born in Seoul, South Korea and raised in the U.S.; Speaker 2 (L1 = Korean, L2 = Spanish) was a female born in the U.S. and raised in Santiago, Chile; Speaker 3 (L1 = Siraeki, L2 = Sindhi) was a male from Sindh, Pakistan; Speaker 4 (L1 = Spanish) was a male from Granada, Spain; and Speaker 5 (L1 = Shanghainese, L2 = Mandarin) was a female from Shanghai, China.

The 24 listeners who provided judgments on the stimuli were native speakers of Korean who reported no cognitive or auditory impairments. They ranged in age from 18 to 64 and were all familiar with the standard *hangeul* orthography.

2.1.4 Procedure

Subjects who participated in the perception experiment were told that they would be taking a test in which they would hear and write down in *hangeul* a series of Korean syllables that had been modified in a speech synthesis program (and that therefore might sound odd to them). The test lasted approximately 15 minutes, and all subjects listened to the 120 syllables in a different random order on a laptop computer over headphones with an inter-stimulus interval of approximately 4.5 seconds. As they heard each syllable only once and were not allowed a break during the experiment, subjects were instructed to write down only their first impression of the syllable and not to worry about transcribing coda consonants, even if a syllable sounded like it was closed.

2.2 Results

In the tables presented below, the labels running vertically correspond to the stimulus category[2], while the labels running horizontally correspond to the laryngeal category of subjects' perceptual judgments – either LEN(IS), FOR(TIS), or ASP(IRATED). All data are presented in terms of the percentage of total responses to a particular stimulus category, and the most common response to each stimulus is in bold. Responses which were not clearly lenis, fortis, or aspirated judgments (e.g. identifications of a stimulus as a sonorant) are excluded; thus, not all rows in the tables below add up to 100%.

---

[2] In the transcription conventions used here, /p/ represents a lenis plosive, /p*/ a fortis plosive, and /p^h/ an aspirated plosive.

2.2.1 Laryngeal categorization of Korean stimuli

As seen in Table 1, subjects are very good at identifying Korean word-initial obstruents correctly.

Table 1: Laryngeal categorization of word-initial Korean obstruents (% of responses).

		Perceptual judgments		
		LEN	FOR	ASP
Korean stimulus	p	**73**	6	21
	p*	33	**63**	4
	p^h	0	0	**100**
	t	**67**	0	33
	t*	17	**83**	0
	t^h	6	0	**94**
	s	**98**	0	0

		Perceptual judgments		
		LEN	FOR	ASP
Korean stimulus	tʃ	**98**	0	2
	tʃ*	19	**73**	8
	tʃ^h	11	0	**89**
	k	**68**	7	25
	k*	8	**88**	4
	k^h	10	0	**90**
	s*	33	**67**	0

The most misidentifications are of lenis plosives as aspirated plosives (bilabial: 21%, alveolar: 33%, velar: 25%), which may be attributed to the overlap in aspiration between these two series (cf. M. Kim 2004): both are aspirated, although the aspirated series tends to be more heavily aspirated than the lenis series. On the whole, however, perception is quite accurate, a fact that is reflected visually in the perfect diagonal of most common responses for each place of articulation.

In the perception of originally intervocalic consonants that now appear to be initial (due to the fact that there is no longer any material preceding them), there is much more divergence from the diagonal of "correct" answers seen in Table 1.

Table 2: Laryngeal categorization of intervocalic Korean obstruents (% of responses).

		Perceptual judgments		
		LEN	FOR	ASP
Korean stimulus	p	36	**62**	2
	p*	21	**75**	4
	p^h	35	0	**65**
	t	39	**59**	2
	t*	19	**79**	2
	t^h	**67**	0	33
	s	**96**	2	0

		Perceptual judgments		
		LEN	FOR	ASP
Korean stimulus	tʃ	30	**64**	6
	tʃ*	8	**79**	13
	tʃ^h	38	4	**58**
	k	**58**	25	0
	k*	18	**75**	7
	k^h	**77**	2	21
	s*	10	**90**	0

This divergence shows some clear patterns (cf. Table 2). First, intervocalic lenis obstruents, which are allophonically voiced, are generally perceived as fortis (bilabial: 62%, alveolar: 59%, post-alveolar: 64%). This is not so in the case of velars, but even here the lenis stops are identified as fortis 25% of the time.

Second, intervocalic aspirated plosives and the aspirated affricate are not perceived with nearly the same amount of veridicality as the initial ones seen in Table 1. For alveolars and velars, the majority of responses are in favor of lenis (alveolar: 67%, velar: 77%), and there are large numbers of responses in this direction for the bilabials (36%) and the post-alveolars (30%) as well.

These perceptual patterns may again be understood in terms of aspiration. When intervocalic lenis obstruents become voiced, they lose the aspiration they have word-initially. Thus, when originally intervocalic lenis obstruents are made to be word-initial, a critical cue to their identity in word-initial position – aspiration – is suddenly missing; consequently, they are identified as fortis. This sort of aspiration reduction may also underlie the perceptual confusion of aspirated and lenis obstruents seen in Table 2. While aspirated obstruents are not voiced intervocalically like lenis consonants, they are indeed significantly less aspirated intervocalically than word-initially. It follows that this relatively shortened aspiration interval may be perceived as closer to the lighter aspiration of the lenis obstruents than the heavier aspiration of the aspirated obstruents (cf. §3.2).

2.2.2 Laryngeal categorization of Sindhi stimuli

Subjects' perception of Sindhi consonants is very consistent, as seen in the similar pattern of bolded cells at each place of articulation in Table 3.

Table 3: Laryngeal categorization of Sindhi obstruents (% of responses).

		Perceptual judgments					Perceptual judgments		
		LEN	FOR	ASP			LEN	FOR	ASP
Sindhi stimulus	p	11	**83**	4	Sindhi stimulus	tʃ	31	**46**	23
	p^h	13	0	**79**		tʃ^h	6	1	**93**
	b	**88**	12	0		dʒ	**96**	3	0
	b^ɦ	**82**	4	13		dʒ^ɦ	**69**	0	7
	t	15	**75**	4		k	25	**69**	4
	t^h	13	0	**83**		k^h	4	0	**96**
	d	**85**	15	0		g	**75**	25	0
	d^ɦ	**79**	0	3		g^ɦ	**92**	4	4
	z	**83**	17	0		ɗ	**33**	8	0

First, voiceless aspirated obstruents are generally perceived as aspirated (bilabial: 79%, dental: 83%, palatal: 93%, velar: 96%), while voiceless unaspirated obstruents are generally perceived as fortis (bilabial: 83%, dental: 75%, palatal: 46%, velar: 69%). Second, all voiced obstruents, both aspirated and unaspirated, are perceived as lenis, the latter result standing in contrast to the perception of the allophonically voiced lenis obstruents of Korean, which, as seen in Table 2, are most often perceived as fortis and occasionally as lenis (with the exception of the velar, for which the pattern goes in the other direction). Table 3 shows voiced unaspirated obstruents being perceived most often as lenis (bilabial: 88%, alveolar: 85%, palatal: 96%, velar: 75%), with a minority of fortis identifications for plosives at all places of articulation. In short, even though in both cases subjects are hearing voiced unaspirated obstruents, the perceptual patterns in Table 2 and Table 3 are essentially the reverse of each other. These results are most likely attributable to differences in strength of voicing as well as differences in closure duration (cf. §4).

The general pattern for Sindhi plosives also applies to the implosives, which are more often perceived as lenis than as fortis. However, the most common perceptual judgments for the implosive stimuli are sonorant identifications. Again, this implies that subjects are attuned to the robust voicing in these stimuli. In this case they appear to interpret the voicing as indicative of a sonorant (all of which are strongly voiced in Korean).

Finally, the perception of the voiced alveolar fricative is somewhat surprising. It is not perceived as either of the voiceless fricatives, but instead as the lenis post-alveolar affricate, suggesting that listeners are sensitive either to the vocal fold vibration that occurs over the segment's duration or to its relatively lower center of gravity. The latter cue in particular lines up quite well with features of the affricate, as the post-alveolar frication of the affricate also has a center of gravity that is lower than that characteristic of the voiceless alveolar fricative.

2.2.3 Laryngeal categorization of Spanish stimuli

The perception of Spanish consonants shows some patterns similar to the perception of Sindhi consonants.

Table 4: Laryngeal categorization of Spanish obstruents (% of responses).

		Perceptual judgments		
		LEN	FOR	ASP
Sp. stimulus	p	**52**	44	0
	b	**96**	4	0
	t	**56**	40	2
	d	**100**	0	0
	s	37	**63**	0

		Perceptual judgments		
		LEN	FOR	ASP
Sp. stimulus	tʃ	**79**	4	17
	dʒ	**92**	0	0
	k	**58**	42	0
	g	**92**	8	0

Like Table 3, Table 4 shows voiced obstruents generally being perceived as lenis (bilabial: 96%, dental: 100%, palato-alveolar: 92%, velar: 92%). On the other hand, the perception of voiceless plosives is split between lenis and fortis. In contrast to the perception of Sindhi voiceless plosives, Spanish voiceless plosives are perceived as fortis less than half of the time; perception instead slightly favors lenis for the bilabial (52%) and velar (58%) and heavily favors lenis for the palato-alveolar affricate (79%). The voiceless alveolar fricative, however, is perceived most often as fortis (63%).

Since the average voice onset time (VOT) of the Spanish voiceless plosives is only slightly longer than that of the Korean fortis plosives (Spanish: 33 ms, Korean: 29 ms, cf. §3.2), it appears that differences in fundamental frequency (f_0) onset are probably responsible for the ambivalent pattern seen here: the average f_0 onset of the Spanish voiceless plosive stimuli is nearly 80 Hz lower than that of the originally word-initial Korean fortis plosive stimuli. This difference in f_0 onset is likely to result in some lenis judgments because it mimics a similar difference in f_0 onset between Korean lenis and fortis plosives.

2.2.4 Laryngeal categorization of Shanghainese stimuli

The perception of Shanghainese voiced obstruents is consistent with that of Spanish and Sindhi voiced obstruents.

Table 5: Laryngeal categorization of Shanghainese voiced obstruents (% of responses).

		Perceptual judgments		
		LEN	FOR	ASP
Shang. stimulus	b	**100**	0	0
	d	**100**	0	0
	g	**83**	13	4

Similar to Tables 3 and 4, Table 5 shows that subjects generally identify voiced obstruents as lenis. In fact, they are completely consistent in identifying bilabial and alveolar voiced plosives as lenis, with 100% of judgments going in this direction. Judgments on the velars are also overwhelmingly lenis judgments, with 83% going in this direction.

2.3 Summary

In general, Korean word-initial obstruents and originally intervocalic fortis obstruents are perceived accurately, but originally intervocalic aspirated obstruents are often perceived as lenis and originally intervocalic (and thus allophonically voiced) lenis obstruents are usually perceived as fortis. Sindhi voiceless aspirated obstruents are perceived as aspirated; voiceless unaspirated obstruents as fortis; and all voiced categories as lenis, including the voiced fricative and the implosive. Unlike Sindhi, Spanish voiceless obstruents are perceived slightly more often as lenis, but similar to Sindhi, Spanish voiced obstruents as well as Shanghainese voiced obstruents are perceived as lenis.

3 Experiment 2: acoustic analysis

Experiment 1 established relationships of perceptual similarity between Korean laryngeal categories and those of Sindhi, Spanish, and Shanghainese, but did not confirm what underlies these patterns. What features of the acoustic signal are affecting the perceptual patterns seen in Experiment 1? To answer this question, acoustic analysis was performed on all stimuli, and these acoustic measures were cross-correlated with the judgment data from Experiment 1.

As noted in §1, the Korean lenis, fortis, and aspirated consonants have been observed to differ from each other acoustically in a number of ways. To examine acoustic correlates of the perceptual patterns observed in Experiment 1, the stimuli used were analyzed with respect to the two most well-documented differences – VOT and f_0 onset – as well as first formant onset, voice quality (as measured by the tilt or slope of the spectrum), and intensity buildup.

3.1 Methods

All acoustic measurements were taken in Praat 4.2.17. VOT was measured from the beginning of the consonant burst (or, in the case of the fricatives, the end of noise concentrated in the high-frequency range) to the onset of glottal vibration; f_0 onset was measured across the first 25 ms of the vowel; first formant (F1) onset was measured at the beginning of the first glottal cycle; spectral tilt was measured via the amplitude difference between the first and second harmonics (H1-H2) over an average spectrum of the first four glottal cycles; and intensity buildup (ΔInt) was determined as the average of intensity differences between adjacent glottal cycles for the first four glottal cycles.

3.2 Results

Tables 6 and 7 present the acoustic data for the word-initial and originally intervocalic Korean stimuli, respectively, averaging over both speakers and all places of articulation.

Table 6: Acoustic measures for Korean word-initial stimuli.

		Acoustic measures				
		VOT (ms)	f_0 (Hz)	F1 (Hz)	H1-H2 (dB)	ΔInt (dB/cyc)
Laryngeal category	LEN	66	146	718	2.5	1.3
	FOR	29	181	631	-2.9	0.8
	ASP	112	182	801	4.1	0.8

Table 7: Acoustic measures for Korean intervocalic stimuli.

		Acoustic measures				
		VOT (ms)	f_0 (Hz)	F1 (Hz)	H1-H2 (dB)	ΔInt (dB/cyc)
Laryngeal category	LEN	-81	154	562	-3.0	0.7
	FOR	31	155	620	-3.8	0.6
	ASP	77	151	727	1.6	0.8

For the word-initial stimuli, VOT is very short for fortis, longer for lenis, and very long for aspirated. The f_0 onset of fortis and aspirated is similar, and for both on the order of 40 Hz above that of lenis. F1 starts lowest for fortis, higher for lenis, and highest for aspirated. As for spectral tilt, H1-H2 is negative for fortis, but positive for lenis and even higher for aspirated. Finally, average intensity buildup is highest for lenis and similar between fortis and aspirated.

The data for the originally intervocalic Korean stimuli look different in a number of ways. The lenis series has voicing prior to closure release, resulting in negative VOT measures. In addition, F1 starts lowest for lenis instead of fortis; H1-H2 is now negative for lenis; and the aspirated series is much less aspirated (nearly 40 ms less). Finally, f_0 differences between the three series are neutralized.

Sindhi plosives are similar to Korean plosives in a few ways (cf. Table 8). VOT of voiceless unaspirated is about the same as for Korean fortis, while VOT of voiceless aspirated is even longer than for Korean aspirated. VOT of voiced unaspirated and aspirated is more strongly negative than for Korean medial lenis, and VOT of the implosives is about half this length. With respect to f_0, f_0 onset is 10-15 Hz lower for voiced unaspirated than for voiceless plosives, and 10-20 Hz lower still for voiced aspirated plosives and implosives. F1 starts lowest for the voiced plosives and starts highest for voiceless aspirated. In terms of spectral tilt, H1-H2 is negative for the voiceless plosives, even more negative for implosives, about zero for voiced unaspirated, and positive for voiced aspirated. Finally, intensity buildup is similar across the plosives and significantly lower for the implosives.

Table 8: Acoustic measures for Sindhi stimuli.

		Acoustic measures				
		VOT (ms)	f_0 (Hz)	F1 (Hz)	H1-H2 (dB)	ΔInt (dB/cyc)
Laryngeal category	VCLS UNASP	34	177	548	-3.2	1.2
	VCLS ASP	140	182	710	-1.3	1.5
	VCD UNASP	-130	167	451	0.2	1.5
	VCD ASP	-124	148	507	3.6	1.5
	IMPLOSIVE	-62	155	587	-6.6	0.8

In the case of the Spanish stimuli (cf. Table 9), voiceless plosives are similar to fortis in VOT, while voiced plosives are, like Sindhi, prevoiced sooner than Korean medial lenis plosives. On the other hand, the f_0 onset of both voiced and voiceless is very low (which, as mentioned earlier, probably encouraged the labeling of voiceless plosives as lenis).

Table 9: Acoustic measures for Spanish stimuli.

		Acoustic measures				
		VOT (ms)	f_0 (Hz)	F1 (Hz)	H1-H2 (dB)	ΔInt (dB/cyc)
Laryngeal category	VCLS UNASP	33	105	491	-5.0	0.9
	VCD UNASP	-137	102	414	-4.6	1.3

Finally, a close look at the Shanghainese "voiced" stimuli reveals that there is no actual vocal fold vibration during closure, but positive VOT approximating lenis. H1-H2 is also highly positive, which is again similar to lenis.

Table 10: Acoustic measures for Shanghainese stimuli.

		Acoustic measures				
		VOT (ms)	f_0 (Hz)	F1 (Hz)	H1-H2 (dB)	ΔInt (dB/cyc)
Laryngeal category	VCD UNASP	61	180	364	2.9	0.7

A linear regression analysis shows that subjects' perceptual judgments in Experiment 1 are significantly correlated with several properties of the acoustic signal. The percentage of lenis identifications is correlated with VOT (r = -.400, p < .0005), f_0 onset (r = -.342, p < .0005), and spectral tilt (r = .327, p = .002), but not with F1 onset[3] (r = -.057, p > .5) or intensity buildup (r = .071, p > .3). The percentage of aspirated identifications is correlated with VOT (r = .423, p < .0005), F1 onset (r = .229, p = .006), and spectral tilt (r = .244, p = .001), but not with f_0 onset (r = .039, p > .5) or intensity buildup (r = .113, p > .1).

The linear regression analysis also shows that percentage of fortis identifications is correlated with f_0 onset (r = .423, p < .0005), spectral tilt (r = .244, p = .001), and intensity buildup (r = -.204, p = .014), but not with F1 onset (r = -.079, p > .3) or VOT (r = .057, p > .5). Note, however, that the latter VOT result is not valid because the data violate a fundamental assumption of linear regression – namely, that the dependent and independent variables are linearly related. In this case, it is clear that percentage of fortis identifications does not change linearly with VOT: only zero/short-lag VOT (neither strongly negative nor positive) results in a robust fortis percept. In other words, the general shape of the data distribution is not a line, but a spike[4]. Consequently, the statistical analysis of the correlation between VOT and fortis identifications must be refined.

The spiky shape of this distribution is better fit by a quadratic or cubic model than a linear model (cf. Figure 1). While a linear model accounts for virtually none of the variance in fortis identifications, a quadratic model accounts for about one-third of the variance (r^2 = .333, F(2, 117) = 29.192, p < .0005), and a cubic model does even better (r^2 = .351, F(3, 116) = 20.900, p < .0005).

---

[3] It is possible that F1 onset is indeed significantly correlated with lenis identification (as it is with aspirated identification) but that this correlation is being obscured by the fact that not all the lenis stimuli had low vowels (i.e. high F1 targets). It is difficult to say how the statistics would have worked out had the vowel quality in these stimuli been completely uniform.

[4] To clarify, the data are distributed such that the function falls to zero/near-zero on either side of the spike (cf. Figure 1). The distribution is not sigmoidal. In the latter case, it would have made sense to perform a logistic regression, but the particular nature of this distribution calls for an analysis that assumes a many-to-one distribution.

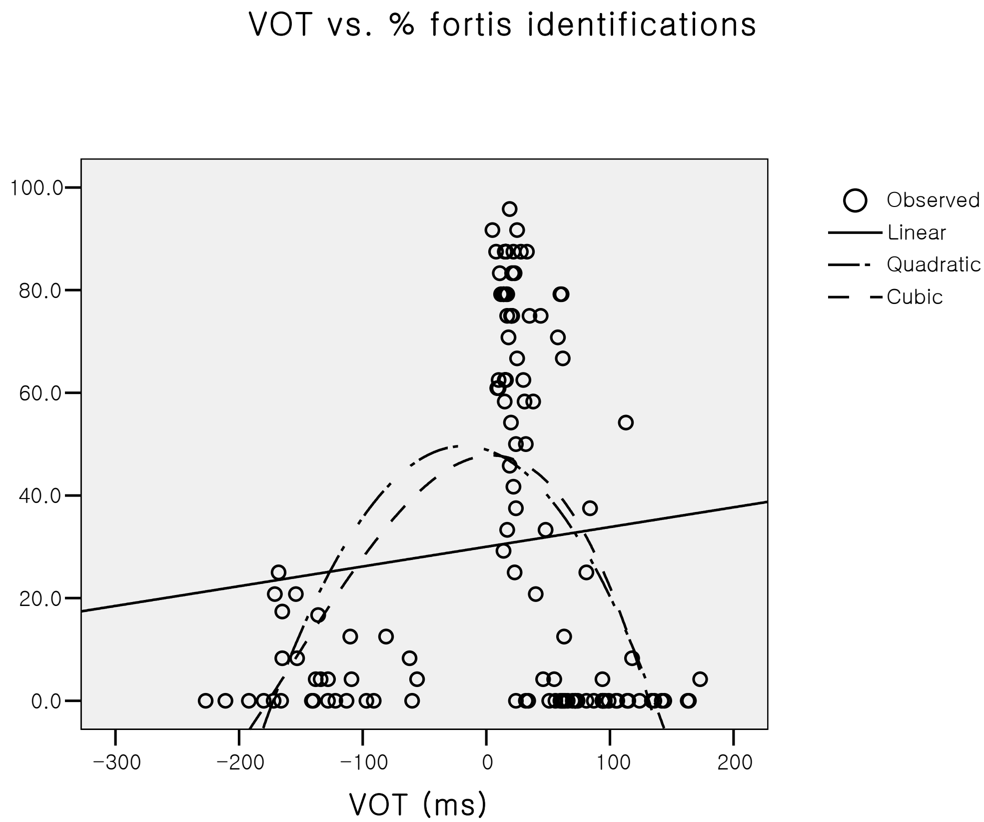

Figure 1: Scatterplot of VOT vs. % fortis identifications with various best-fit models.

In short, it seems that VOT and spectral tilt play a significant role in the perception of all three laryngeal categories, and that f_0 onset is also relevant in the perception of the lenis and fortis categories.

4 Discussion

The findings of this study have implications both for a map of cross-linguistic perceptual similarity as well as for the analysis of Korean laryngeal contrast. With regard to cross-linguistic perceptual similarity, it was observed that (i) Sindhi voiced obstruents (unaspirated and aspirated plosives, implosives, and sibilant /z/), Spanish voiced obstruents, and Shanghainese voiced obstruents are all perceptually similar to Korean lenis obstruents, (ii) Sindhi and Spanish voiceless unaspirated obstruents are perceptually similar to Korean fortis obstruents, (iii) Sindhi voiceless aspirated obstruents are perceptually similar to Korean aspirated obstruents, and (iv) Spanish /s/ is perceptually similar to Korean fortis /s*/.

In this way the data are superficially consistent with an interpretation of the Korean laryngeal contrast along the lines of Kim and Duanmu (2004), who use the f_0 difference between lenis and fortis/aspirated along with other arguments to group the Korean contrast with the familiar voicing and aspiration contrasts found in other languages of the world. Since both Sindhi and Spanish voiced consonants are identified as Korean lenis – in spite of the acoustic data found in Experiment 2, which indicate that, other than voicing/VOT, Sindhi voiced unaspirated is actually similar to Korean fortis – it appears that Korean speakers are attending to vocal fold vibration rather closely. Such a fact is easily accounted for by positing that Korean contains a phonologically voiced category.

However, acoustic data paradoxically do not support positing a phonologically voiced laryngeal category, since the voicing of the Korean voiced lenis obstruents is actually quite weak. Remember that in §2.2.2 it was observed that Korean speakers do not perceive all voiced unaspirated obstruents in the same way: Korean's voiced unaspirated (lenis) plosives are perceived as fortis, while Sindhi's voiced unaspirated plosives are perceived as lenis. As alluded to above, acoustic explanations for this perceptual difference can be found in the phonetic quality of these phones in the two languages. Sindhi's phonologically voiced plosives are more robustly voiced (even slightly pre-nasalized) as well as longer than Korean's allophonically voiced plosives. In fact, the average duration difference between Korean's voiced plosives and Sindhi's voiced plosives is quite substantial – nearly 50 ms (cf. Tables 7 and 8). These acoustic differences can be seen in Figure 2, which compares voicing in Korean [da] extracted from [tada] 'close (it)' with Sindhi [da] in [darũŋ] 'gunpowder'. In Korean [da], voicing is barely visible, starts much closer to the stop burst, and fades quickly, while in Sindhi [da], voicing is very strong, starts much earlier, and lasts throughout the period of closure.

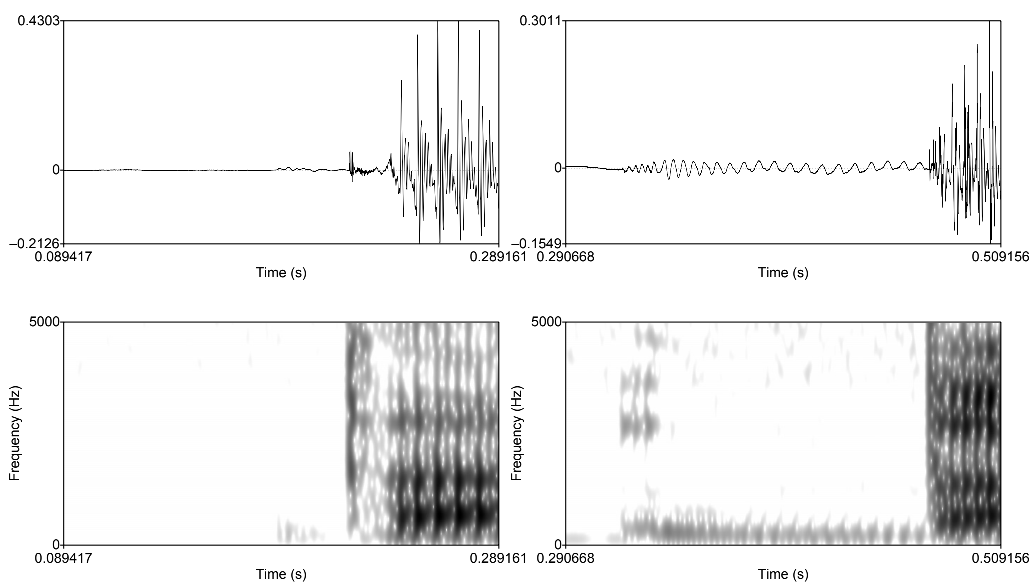

Figure 2: Robustness of voicing in Korean [da] (L) vs. Sindhi [da] (R).

Due to either the relatively weak quality of the Korean voicing, its relatively short duration, or a combination of both of these factors, Korean voiced lenis plosives are, interestingly enough, not perceived as lenis by Korean speakers. This finding weakens Kim and Duanmu's (2004) claim that the Korean lenis category is just another voiced category[5]. If that were the case, we would expect that, at the very least, the voiced allophones in Korean would be perceived as lenis by Korean speakers. The fact that they are not perceived this way suggests that even though Korean speakers generally label voiced obstruents from other languages as lenis, they do so not so much because other languages' voiced categories and Korean's lenis category are very perceptually close, but more so because other languages' voiced categories are *closer* to the Korean lenis category than to the fortis and aspirated categories. In this sense, this labeling pattern may constitute the "lesser of two evils" rather than the result of genuinely close perceptual similarity between categories, which may also account for why some of the judgments in Experiment 1 seem somewhat ambivalent instead of approaching 100% consistency. Thus, Korean voicing seems to be more a case of phonetic voicing than phonological voicing.

Moreover, the perceptual data in Experiment 1 point out the inadequacy of simply calling the lenis series "voiced", since a significant cue to word-initial lenis is aspiration (i.e. a VOT lag), an unusual property for a voiced category. Voiced consonants from other languages are indeed perceived as lenis most of the time, but when there is no VOT lag, voiced plosives and implosives are often perceived as fortis (cf. the minority responses to voiced Sindhi obstruents in Table 3 and the majority responses to allophonically voiced Korean lenis obstruents in Table 2). Note that these perceptual judgments are made even in the face of conflicting cues: allophonically voiced Korean lenis obstruents are perceived as fortis even though their duration is much shorter than that typical of fortis obstruents. Finally, not only do Korean medial lenis plosives (having negative VOT) result in a fortis percept, Korean medial aspirated plosives (having positive VOT) are often misidentified as lenis. These facts suggest that a VOT lag is an important part of the phonetic implementation of word-initial lenis consonants in Korean. In contrast, voiced consonants cross-linguistically are characterized precisely by the lack of such a lag (or, for that matter, a "negative lag" – vocal fold vibration prior to the release of closure). It seems, then, that in spite of the

---

[5] Other problems with Kim and Duanmu's arguments are more theoretical in nature. For instance, they contend that adding the feature [tense] to the universal feature pool in order to be able to specify all three voiceless series in Korean is problematic for phonological theory because the coexistence of [voice], [spread glottis], and [tense] overgenerates unattested sounds such as /b^h*/. However, the coexistence of these features does not necessarily have to overgenerate such sounds if certain features or feature values are simply assumed to be incompatible on phonetic grounds. This sort of feature incompatibility must be assumed in any case to account for why, e.g., there is no voiced glottal stop even though both [constricted glottis] and [voice] are taken to be standard phonological features: it is physically (and logically) impossible to fully close the glottis and yet open it enough for voicing at the same time.

fact that the voiced obstruents of other languages are perceptually closest to lenis obstruents for Korean speakers, the Korean lenis category remains quite different in nature from the prototypical voiced category.

Nonetheless, the basic typological thrust of Kim and Duanmu's argument is a valuable direction for future research. Given that there is some degree of perceptual similarity between Korean lenis obstruents and other languages' voiced obstruents, it is interesting to consider what sort of more general category these laryngeal types may be considered to exemplify.

5 Conclusion

This study reexamined the typologically unusual three-way laryngeal contrast in Korean among lenis, fortis, and aspirated voiceless obstruents in light of a recent proposal by Kim and Duanmu (2004) that the contrast is really among voiced, voiceless unaspirated, and voiceless aspirated categories. The results of a cross-linguistic perception experiment conducted with 24 native Korean speakers are superficially consistent with Kim and Duanmu's analysis: Korean speakers are attuned to voicing in initial position and identify voiced obstruents from other languages as lenis. While this could be accounted for by a Korean laryngeal system containing a phonologically voiced category, two acoustic facts argue against positing phonological voicing in Korean. First, the fact that Korean lenis obstruents appear to be only weakly voiced even in an intervocalic environment favoring voicing suggests that the nature of voicing in Korean is phonetic, not phonological. Second, an important factor in Korean speakers' perception of word-initial lenis obstruents is a feature atypical of voiced obstruents – namely, aspiration. If the lenis series were to be analyzed as voiced, it would thus differ in rather fundamental ways from the phonetic character of other languages' voiced plosives. It seems, then, that "voiced" is simply a box into which the Korean laryngeal contrast fails to fit.

References

Boersma, P., & D. Weenink. 2004. Praat: doing phonetics by computer (Version 4.2.17). http://www.praat.org.

Chang, C. B. 2007. Perspectives on the Korean laryngeal contrast from cross-linguistic perceptual similarity. In *LSO Working Papers in Linguistics* 5, ed. by Rebecca Shields. Madison, WI: UW-Madison Linguistics Student Organization.

Cho, T., & P. Keating. 2001. Articulatory and acoustic studies of domain-initial strengthening in Korean. *J. Phonetics* 29(2).155–190.

Cho, T., S.-A. Jun, & P. Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *J. Phonetics* 30(2).193–228.

Choi, H. 2002. Acoustic cues for the Korean stop contrast – dialectal variation. *ZAS Papers in Linguistics* 28.1–12.

Dart, S. 1987. An aerodynamic study of Korean stop consonants: measurements and modeling. *J. Acoust. Soc. Am.* 81(1).138–147.

Han, J.-I. 1996. The phonetics and phonology of “tense” and “plain” consonants in Korean. Ph.D. dissertation, Cornell University.

Han, M. S., & R. S. Weitzman. 1970. Acoustic features of Korean /P, T, K/, /p, t, k/, and /p^h, t^h, k^h/. *Phonetica* 22.112–128.

Han, N. 1998. A comparative acoustic study of Korean by native Korean children and Korean-American children. M.A. thesis, University of California, Los Angeles.

Hardcastle, W. J. 1973. Some observations on the tense-lax distinction in initial stops in Korean. *J. Phonetics* 1.263–272.

Hirose, H., C. Y. Lee, & T. Ushijima. 1974. Laryngeal control in Korean stop production. *J. Phonetics* 2.145–152.

Kagaya, R. 1974. A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *J. Phonetics* 2.161–180.

Kim, C.-W. 1965. On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word* 21.339–359.

Kim, C.-W. 1970. A theory of aspiration. *Phonetica* 21.107–116.

Kim, M. 2004. Correlation between VOT and F0 in the perception of Korean stops and affricates. In *INTERSPEECH 2004*, 49–52.

Kim, M.-R. 1994. Acoustic characteristics of Korean stops and perception of English stop consonants. Ph.D. dissertation, University of Wisconsin, Madison.

Kim, M.-R., & S. Duanmu. 2004. Tense and lax stops in Korean. *J. East Asian Linguistics* 13.59–104.

Lee, H. B. 1999. Korean. In *Handbook of the International Phonetic Association*, 120–123. Cambridge: Cambridge University Press.

Lisker, L., & A. Abramson. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20(3):384–422.

Park, H. 2002. The time courses of F1 and F2 and a descriptor of phonation types. *Eoneohag* [Linguistic Sciences] 33.87–108.

Silva, D. 1992. The phonetics and phonology of stop lenition in Korean. Ph.D. dissertation, Cornell University.