Morpheme structure and co-occurrence restrictions in Korean monosyllabic stems *

Chiyuki Ito
(Tokyo University of Foreign Studies)

Chiyuki, Ito. 2007. Morpheme structure and co-occurrence restrictions in Korean monosyllabic stems. *Studies in Phonetics, Phonology and Morphology 0.0 00-000.* This paper reports the results of a study of the statistical distribution of various phonological contrasts in Korean monosyllabic stems. The data is composed of 1,419 native (not Sino-Korean) roots drawn from an authoritative dictionary. Major findings are: (1) Segmental frequencies in onset, nucleus, and coda coincides with a cross-linguistic markedness hierarchy such that less marked segments/classes are more attested than marked ones in the lexicon. Verbs (verbal stems) show a more complex coda inventory than nouns, which is due to the Base-Identity effect (Kenstowicz 1996). (2) As in many genetically unrelated languages, Korean has an OCP-Place constraint that is statistically significant (Greenberg 1950, McCarthy 1986, 1988, 1994, Mester 1986, Padgett 1995, Frisch et al. 2004, Kawahara et al. 2006, Coetzee and Pater 2006); (3) Korean has a gradient complexity constraint that restricts stems composed of more than one segment drawn from the following classes: aspirated/tense consonants, diphthongs, complex codas (Anttila 2007. cf. MacEachern 1997); (4) The distribution of aspiration is predictable. It occupies the coda position as a rule, but moves to the onset position when the coda is either a segment which cannot bear aspiration or -k. k^h is highly underrepresented in the Korean lexicon; (5) To place the extreme underrepresentation of Korean aspirated velar stops in perspective, we conducted a typological survey of language frequency for aspiration contrasts based on the data in Maddieson (1984). The implicational hierarchy is Coronal > Labial, Velar > Uvular for place and Stop > Affricate > Fricative for manner. The extreme underrepresentation of k^h in Korean is not due to the cross-linguistic dispreference for k^h but probably results from potential confusion with a velar/uvular fricative. Thus the findings of this paper not only contribute new data to the literature on Korean phonetics/phonology, but also reveal a cross-linguistic hierarchy in the distribution of aspiration with respect to oral place. **(Tokyo University of Foreign Studies)**

Keywords: co-occurrence restriction, OCP-Place constraint, gradience, *COMPLEX, aspiration, typology

1. Introduction

A number of researchers have examined Korean both phonetically and phonologically, and have revealed its unique properties from various perspectives. However as far as we know, there is no single work that analyzes the Korean phonotactics in detail, by providing a sufficient amount of data. This paper reports the major results of a study of the statistical distribution of various phonological contrasts in Korean monosyllabic verbal/nominal stems, based on the data from *Pyojun kugeo taesajeon* [Standard Korean dictionary] (Kungnip kugeo yeonguwon (Ed.),

---

* This paper was originally presented at the 15th Manchester Phonology Meeting. I thank Adam Albright, Daniel Silverman, Michael Kenstowicz, Stuart Davis, as well as three anonymous reviewers for helpful comments.

1999, Seoul: Tusandong'a). The total number of our data is 1,419 stems (664 for verbs, and 755 for nouns), and all entries are native or highly nativised words (not Sino-Korean).

The rest of this paper is organized as follows. In Section 2, segmental frequencies in onset, nucleus, and coda are examined, whereby the correlation between the segmental distribution and a cross-linguistic markedness hierarchy is shown. In Section 3, two co-occurrence restrictions are pointed out. First, as in many genetically unrelated languages, Korean has an OCP-Place constraint that is statistically significant (Greenberg 1950, McCarthy 1986, 1988, 1994, Mester 1986, Padgett 1995, Frisch et al. 2004, Kawahara et al. 2006, Coetzee and Pater 2006). Second, Korean has a gradient complexity constraint that restricts stems composed of more than one segment drawn from the following classes: aspirated/tense consonants, diphthongs, complex codas (cf. MacEachern 1997). In Section 4, the positional predictability of aspiration in Korean monosyllabic stems is proposed, along with an analysis within an Optimality Theory framework (Prince and Smolensky 1993). In Section 5, a typological survey of language frequency for aspiration contrasts is conducted based on the data in Maddieson (1984), by which we examine the striking underrepresentation of k^h in Korean. Section 6 is the conclusion. All the statistical frequency reported in this paper is based on type-frequency, not token-frequency.

2. Korean sound system and segmental frequency [1]

2.1 Consonants

Table 1 shows the inventory of Korean consonants. As is well known, Korean obstruents have a three-way laryngeal contrast (lax, aspirated, and tense), such as p/p^h/p', t/t^h/t'. The categorization of s, which lacks the opposition between the lax series and the aspirated series, has been controversial (Kagaya 1974, Iverson 1983, Jun 1993, Park 1999, Cho et al. 2002). In Korean, ŋ does not appear in syllable-initial position. Liquid, which appears as [r] in syllable-initial position and as [l] in syllable-final position, does not appear in word-initial position except for recent loanwords.

[1] The reviewer raised a relevant question whether the context for the statement of phonotactic constraints is best formulated in terms of syllable positions such as onset and coda, or as an initial-medial-final position in the stem, regardless of syllable structure. It is not easy to distinguish between these two possibilities. Many of the restrictions seem to correspond to natural cross-linguistics restrictions on syllable structure, such as the distributional bias between obstruent and sonorant discussed in 2.1. On the other hand, given the distribution of liquid which does not appear in word-initial position of native lexicon but can appear as an onset in word-medial position, it is not sufficient to rely on only syllable structure. Determining whether syllable structure or stem position is the relevant context for the statement of any given phonotactic constraint is a task for future research.

Table 1. Consonants

Lax	p	t	c [tʃ]		k	
Aspirated	p^h	t^h	c^h	s	k^h	h
Tense	p’	t’	c’	s’	k’	
Nasal	m	n			ŋ	
Liquid		r/l				

Table 2 shows the segmental frequency of onsets in Korean monosyllabic stems. Roughly speaking, Korean onsets are divided into two frequency groups: based on the median split (76.3 when excluding Ø, 78.8 when including Ø), high-frequency onsets (k, s, p, m, c, t, n, and Ø) are lax consonants, sonorants, s, and Ø, whereas low-frequency onsets (c^h, k’, s’, t^h, c’, h, p^h, p’, k^h) are aspirated/tense consonants and h (t’ can belong to both groups depending on the treatment of Ø in the median calculation). The common phonological assumption concerning laryngeal features is that plain voiceless consonants are less marked than voiceless aspirated consonants (for an opposing view about the laryngeal markedness of aspiration, see Vaux and Samuels 2005). A strong bias observed in Table 2 (Lax {p, t, c, k} = 494, Aspirated {p^h, t^h, c^h, k^h} = 167, Tense {p’, t’, c’, k’} = 207) suggests that as in many other languages, the lax (plain) series is unmarked in Korean, whereas the aspirated and tense series are marked.

Table 2. Onset frequency

Onsets	Verbs	Nouns	Totals	Examples
k	71	87	158	ka- ‘go,’ ke ‘crab’
s	50	79	129	sa- ‘buy,’ so ‘cow’
p	45	78	123	po- ‘see,’ pi ‘rain’
m	49	72	121	mɛ- ‘tie,’ me ‘hammer’
c	57	58	115	ca- ‘sleep,’ co ‘millet’
t	42	56	98	tu- ‘put,’ tɛ ‘bamboo’
n	43	46	89	na- ‘come out,’ na ‘I’
t’	37	41	78	t’a- ‘pick,’ t’i ‘belt’
c^h	33	33	66	c^ha- ‘be cold,’ c^ho ‘candle’
k’	40	22	62	k’a- ‘peel,’ k’ɛ ‘sesame’
s’	35	15	50	s’a- ‘wrap,’ s’i ‘seed’
t^h	18	29	47	t^ha- ‘ride,’ t^he ‘frame’
c’	26	16	42	c’a- ‘squeeze,’ c’i ‘tag’
h	16	24	40	ha- ‘do, say,’ hɛ ‘sun’
p^h	18	14	32	p^ha- ‘dig,’ p^hi ‘blood’
p’	15	10	25	p’ɛ- ‘pull out,’ p’jə ‘bone’
k^h	10	12	22	k^hɛ- ‘dig,’ k^hi ‘height’
Ø	59	63	122	o- ‘come,’ i ‘tooth’
Totals	664	755	1419	

On the other hand, Table 3 shows the *Observed*, *Expected*, and *Observed/Expected values* of the stems with onsets derived from each laryngeal category. The left numbers divided by slash marks indicate the Observed value whereas the right numbers indicate the Expected values. Expected values to each Observed value is calculated by assuming that each laryngeal feature can combine with any place of articulation at random. For example, the Expected value of p is calculated as: 180/868 (= probability of labials) × 494/868 (= probability of lax series) × 868 = 102.4 (= probability of lax labial). Thus, the Observed/Expected value of p, which is shown by a parenthesized number, is 123/102.4 = 1.20.

O/E greater than 1 indicates overrepresentation whereas a value of O/E less than 1 indicates underrepresentation (cf. Pierrehumbert 1994). In this paper, we adjust O/E values for underrepresentation by using upper confidence limit statistics in order to investigate the significance of underrepresentation (Mikheev 1997, Albright & Hayes 2002, Albright 2002a, 2002b, in press. cf. http://davidmlane.com/hyperstat/B9168.html). [2] The confidence value α is 0.75. For example, the (corrected) upper confidence limit of t is: O/E + Standard Error × Two-tailed z value + (0.5/E) = 98/126.9 + 0.04 × 1.15 + (0.5/126.9) = 0.82. The adjusted numbers are given in square brackets. Bold numbers show the cases of underrepresentation.

Table 3. Correlation between laryngeal features and place of articulation (Onset)

Onsets	Lax	Aspirated	Tense	Totals
p/pʰ/p’	123/102.4 (1.20)	**32/34.6 (0.92 [0.99])**	**25/42.9 (0.58 [0.68])**	180
t/tʰ/t’	**98/126.9 (0.77 [0.82])**	47/42.9 (1.10)	78/53.2 (1.47)	223
c/cʰ/c’	**115/126.9 (0.91 [0.94])**	66/42.9 (1.54)	**42/53.2 (0.79 [0.86])**	223
k/kʰ/k’	158/137.7 (1.15)	**22/46.6 (0.47 [0.57])**	62/57.7 (1.07)	242
Totals	494	167	207	868

The distribution as a whole is statistically significant ($\chi^2 = 62.53$, $p < 0.001$): Korean laryngeal features tend to correlate with place of articulation.

The Korean tense series are voiceless laryngealized obstruents. Laryngealized sounds have been traditionally assigned to different phonetic categories from ejectives, mainly because a glottal stricture in laryngealized sounds is superimposed on pulmonic egressive airstream and not used as an initiator of the airstream (Maddieson 1984, Ladefoged and Maddieson

[2] I thank Adam Albright for helpful information.

1996). In contrast, phonological analysis has characterized laryngealized sounds, ejectives, and glottal stops by a single feature such as [constricted glottis] (Chomsky and Halle 1968, Halle and Stevens 1971, Lombardi 1994, Clements 2003). As Table 3 shows, in the tense series p'- and c'- are underrepresented whereas t'- and k'- are overrepresented. This distributional difference may be explained by the implicational typology of ejective segments (Greenberg 1970, Maddieson 1984): if a language has ejective affricates it also has ejective stops; if a language has /p'/, it also has /t'/ (dentals or alveolars); if a language has /t'/, it also has /k'/. Although the degree of overrepresentation is higher in t'- than in k'-, the Korean distribution partially supports the cross-linguistic correlation between laryngealized sounds and ejectives. [3]

As seen in Table 3, the distribution of the lax series tends to be the mirror image of the aspirated series: t-, c-, p^h-, k^h- are underrepresented whereas p-, k-, t^h-, c^h- are overrepresented. In particular k- and k^h- show this reverse correlation strikingly. As discussed below, this idiosyncratic distribution has resulted from the stem-level phonotactics.

Coda distribution has similar tendencies as onsets. (1) shows the Korean morphophonemic coda consonants, which neutralize in syllable-final position as in (2).

(1) Morphophonemic coda consonants
a. Simplex: -p, -p^h, -m, -t, -t^h, -c, -c^h, -s, -s', -n, -l, -k, -k^h, -k', -ŋ, -h
b. Complex: -ps, -lp, -lp^h, -lm, -lt^h, -ls, -lk, -lh, -nc, -nh, -ks

(2) Coda neutralizations
a. -p, -p^h -ps, -lp^h, > [-p˺]
b. -m, -lm > [-m]
c. -t, -t^h, -c, -c^h, -s, -s', -h > [-t˺]
d. -k, -k^h, -k', -ks > [-k˺]
e. -n, -nc, -nh > [-n]
f. -l, -lp, -lt^h, -ls, -lh > [-l] (palp- 'step' is often pronounced as [pap˺])
g. -lk > [-k˺] or [-l]

Some codas, which are represented by capitals -P, -T, -S in this paper, alternate between unreleased stops [-p˺], [-t˺] in syllable-final position and [w], [r], Ø in syllable-initial position. In Middle Korean (15-16th C, MK), these codas alternated between [-p], [-t], [-s] in syllable-final position and voiced lenis [β], [r], [z] in syllable-initial position, respectively. (It is unclear whether stop codas were released or not in MK.) The transcription of the coda -S as in (3) is based on this historical background, in that it alternates between [-t˺] and Ø in Contemporary Korean and alternated

[3] The tense series derived from consonant clusters in Middle Korean (sk-, pk-, psk- > k'-; st-, pt-, pst- > t'-; pc- > c'-; sp- > p'-), but the phonetic features of the second/third segments in the consonant clusters may have been similar to those of current tense consonants.

between [-s] and [z] in MK. -lP is included in lenis codas, reflecting the alternation [-l]~[-lw] which is parallel with the alternation in -P ([-p˺]~[w]). These lenis codas only appear in verbal stems.

(3) Lenis codas

a. -P ([-p˺]~[w]): kuP- 'burn' kuP-ko [kup˺k'o] 'burn and' kuP-ə-sə [kuwəsə] 'burn, and so'

b. -T ([-t˺]~[r]): kəT- 'walk' kəT-ko [kət˺k'o] 'walk and' kəT-ə-sə [kərəsə] 'walk, and so'

c. -S ([-t˺]~Ø): ciS- 'make' ciS-ko [cit˺k'o] 'make and' ciS-ə-sə [ciəsə] 'make, and so'

d. -lP ([-l]~[-lw]): səlP- 'be sad' səlP-ko [səlk'o] 'be sad and' səlP-ə-sə [səlwəsə] 'be sad, and so'

Table 4 shows the coda distribution. Many languages restrict the complexity of syllable margins: complex onsets/codas are universally marked as compared to simple onsets/codas (*COMPLEX, Prince and Smolensky 1993). In Korean, complex codas are much rarer than simplex codas (885 vs. 73), which coincides with this syllable typology.

Table 4. Coda frequency

Coda	Verbs	Nouns	Totals	Examples
l	106	137	243	al- ‘know,’ al ‘egg’
m	22	120	142	kam- ‘roll up,’ kam ‘persimmon’
k	31	63	94	cak- ‘be small,’ pak ‘gourd’
ŋ		73	73	t’aŋ ‘ground,’ tɨŋ ‘back’
s	13	55	68	pəs- ‘take off,’ cas ‘pine nuts’
n	4	39	43	an- ‘hug,’ an ‘inside’
p	14	21	35	cap- ‘seize,’ pap ‘boiled rice’
c	24	7	31	nac- ‘be low,’ nac ‘daytime’
t^h	14	15	29	kat^h- ‘be identical,’ pat^h ‘field’
t	20	1	21	tat- ‘close,’ nat ‘grain’
lk	14	6	20	nalk- ‘be old,’ talk ‘chicken’
p^h	9	10	19	kap^h- ‘pay back,’ ap^h ‘front’
c^h	4	13	17	coc^h- ‘follow,’ nac^h ‘face’
P	17		17	təP- ‘be hot,’ nuP- ‘lie down’
S	16		16	naS- ‘get well,’ iS- ‘tie’
lh	14		14	alh- ‘be ill,’ silh- ‘dislike’
k’	12	1	13	nak’- ‘fish,’ pak’ ‘outside’
T	12		12	taT- ‘run,’ kəT- ‘walk’
h	10		10	nah- ‘bear a child,’ tah- ‘touch’
lm	8	2	10	salm- ‘boil,’ salm ‘life’
lp	6		6	palp- ‘step,’ nəlp- ‘be wide’
nh	6		6	manh- ‘be numerous,’ k’ɨnh- ‘cut’
ks		5	5	saks ‘wages,’ moks ‘share’
ls		3	3	ols ‘reward,’ tols ‘one year old birthday’
lt^h	3		3	$halt^h$- ‘lick,’ $holt^h$- ‘scrape off’
nc	2		2	anc- ‘sit,’ ənc- ‘put on’
ps	1	1	2	əps- ‘not have,’ kaps ‘price’
k^h		1	1	$pwək^h$ ‘kitchen’
s’	1		1	is’- ‘exist, have’
lP	1		1	səlP- ‘be sad’
lp^h	1		1	$ɨlp^h$ - ‘recite’
Ø	279	182	461	ca- ‘sleep,’ ca ‘scale’
Totals	664	755	1419	

In Korean, the coda inventory is quite different between verbs and nouns. Including coda-zero (Ø), the total number of distinct morphophonemic codas is 28 for verbs and 20 for nouns. -k^h, -ŋ, -ks and -ls appear only in nouns, whereas lenis codas (-P, -S, -T, -lP), -h, -s’, and many complex codas (-lh, -lp, -nh, -lt^h, -nc, -lp^h) appear only in verbs. The fact that verbs tend to contain more marked/complex codas in Korean is probably due to the morphological difference between verbs and nouns (the Base-Identity effect, Kenstowicz 1996). In Korean, verbs always require an inflection but nouns may appear in an isolation form. That is, verbs often realize the morphophonemic codas faithfully before a vowel-initial suffix, whereas nouns have much more chances to neutralize the morphophonemic codas and hence have more difficulties to maintain them faithfully. In fact, historical coda changes from MK to Contemporary Korean support this hypothesis: nouns have lost many coda distinctions whereas verbs have not. For example in nouns, -h, -mh, -nh, -lh, -lp, -lp^h in MK basically merged with -Ø, -m, -n, -l, -p, -p^h respectively. E.g.) coh > co 'millet,' amh > am 'female,' anh > an 'inside,' alh > al 'egg,' salp > sap 'shovel,' alp^h > ap^h 'front.'

Table 5 shows the correlation between laryngeal features and place of articulation in coda. As in onsets, lax consonants are more attested than aspirated /tense consonants, but the number of stems with tense codas is much smaller: only 14 stems have tense codas, and 13 out of 14 tense codas are -k' (the other is -s'). This bias again could be explained by the implicational hierarchy of ejectives (Maddieson 1984): if a language has only one ejective stop, it is /k'/.

Table 5. Correlation between laryngeal features and place of articulation (Coda)

Codas	Lax	Aspirated	Tense	Totals
p/p^h/p'	**35/37.6** **(0.93 [0.99])**	19/13.7 (1.39)	**0/2.7** **(0.00 [0.19])**	54
t/t^h/t'	**21/34.8** **(0.60 [0.71])**	29/12.7 (2.28)	**0/2.5** **(0.00 [0.20])**	50
c/c^h/c'	**31/33.4** **(0.93 [0.99])**	17/12.2 (1.40)	**0/2.4** **(0.00 [0.21])**	48
k/k^h/k'	94/75.2 (1.25)	**1/27.4** **(0.04 [0.10])**	13/5.4 (2.41)	108
Totals	181	66	13	260

In the aspirated series there is almost no example of -k^h. The existence of -k' in coda may make up for the lack of -k^h (complementary distribution). Still as mentioned above, the avoidance of k^h is also observed in the onset. This suggests that k^h is the least optimal aspirated consonant in Korean, regardless of the position in a stem. In fact, as seen in Table 6 which shows the total of each lax/aspirated consonant in the onset and coda, the percentage of aspirated consonants is lowest in velar stops. We examine this problem in Section 4 and Section 5.

Table 6. Totals of lax/aspirated consonants in onset and coda

Onsets & codas	Lax	Aspirated	Totals	Aspirated (%)
p/p^h	158	51	209	24.40
t/t^h	119	76	195	38.97
c/c^h	146	83	229	36.24
k/k^h	252	23	275	**8.36**
Totals	675	233	908	25.66

There is a different distributional pattern between onset and coda in Korean: onset prefers obstruents whereas coda prefers sonorants. Table 7 shows the ratios of obstruents and sonorants in onset and simplex coda. Here obstruents include k, c, t, p, s, h, k', c', t', p', s', k^h, c^h, t^h, p^h, whereas sonorants include m, n, ŋ, l (lenis codas are excluded). The distribution is significant (χ^2 = 433.56, $p< 0.001$; Fisher's Exact Test (2-Tail): p = 1.05e-96). This tendency coincides with the Sonority Cycle (Clements 1990), according to which the sonority profile of the preferred syllable types rises maximally at the beginning and drops minimally at the end. The avoidance of word-initial r in Korean may also reflect this Sonority Cycle.

Table 7. Sonority and the position in a stem

	Onset	Coda	Totals
Obstruent	1087	339	1426
Sonorant	210	501	711
Totals	1297	840	2137

In sum, the general tendency in phonotactic restrictions observed in Korean monosyllabic stems basically follows a cross-linguistic markedness hierarchy and the syllable typology.

2.2 Vowels

The overall distribution of vowels in Korean monosyllabic stems tends to coincide with a cross-linguistic markedness hierarchy as well. Table 8 shows the vowel (monophthong) inventory (Lee and Ramsey 2000). In actuality, ü and ö are pronounced as diphthongs [wi] and [we] more often. The inventory of diphthongs is ja, jə, jo, ju, je, jɛ, wa, wə, we, and wɛ. (ɨj, spelling pronunciation, is included into /i/ in this paper.)

Table 8. Vowels (monophthongs)

	Front Unrounded	Front Rounded	Central	Back
High	i	ü	ɨ	u
Mid	e	ö	ə	o
Low	ɛ		a	

Table 9 shows the nucleus frequency. As a whole, monophthongs are more frequent than diphthongs (1321 vs. 98); peripheral vowels are more frequent than central vowels. Central mid vowels ə, jə may be the exceptions for the latter tendency, because they are relatively frequently attested. This probably has resulted from some historical reasons, such as the status of ə in a processes of vowel harmony and merger, which we do not examine in detail in this paper. On the other hand, the predominance of /a/ probably reflects its highest sonority: it makes a word more salient, and hence /a/ is chosen as an optimal nucleus.

Table 9. Nucleus frequency

Vowels	Verbs	Nouns	Totals	Examples
a	133	143	276	ka- 'go,' ka 'edge'
i	111	87	198	pi- 'be empty,' pi 'rain'
o	63	134	197	po- 'see,' k^ho 'nose'
ə	71	89	160	sə- 'stand,' cə 'I'
u	67	84	151	cu- 'give,' su 'male'
ɛ	63	62	125	kɛ- 'clear up,' kɛ 'dog'
ɨ	55	28	83	k^hɨ- 'be big,' kɨ 'it'
jə	19	40	59	k^hjə- 'turn on,' hjə 'tongue'
ö/we	30	14	44	k'ö-/k'we- 'tempt,' k'ö/k'we 'wisdom'
ü	24	20	44	cü- 'grasp,' cü 'mouse'
e	19	23	42	me- 'carry,' me 'hammer'
wa	1	10	11	kwal- '(flames) be furious,' hwal 'bow'
ja	3	6	9	jak- 'be wise,' p'jam 'cheek'
wə		7	7	wəl 'sentence,' k'wəŋ 'pheasant,'
je	2	2	4	je- 'go (old style),' je 'ancient times'
wɛ	1	3	4	kwɛl- '(the gold content) be little,' hwɛ 'torch'
ju		2	2	juc^h 'the Four-Stick Game,' hjuŋ 'wound'
ö	1		1	t'ö- 'imply' (clerical error of t'ö- ~ t'we-?)
jo		1	1	jo 'mattress'
we	1		1	k'we- 'thread'
Totals	664	755	1419	

3. Co-occurrence restriction

3.1 OCP-Place constraint

There is a well-known restriction against homorganic consonants in adjacent positions within the verbal root in Arabic, by which the existence of an OCP-Place constraint (Obligatory Contour Principle for Place of Articulation) is motivated (Greenberg 1950, McCarthy 1986, 1988, 1994, Frisch et al. 2004). Similar co-occurrence restrictions are reported in several genetically unrelated languages such as Javanese (Mester 1986), Russian (Padgett 1995), Japanese (Kawahara et al. 2006), and Muna (Coetzee and Pater 2006). In reality, the OCP-Place constraint is observed in Korean as well. Table 10 shows the number of Korean monosyllabic stems which are composed of each onset and simplex coda. For example, the number of stems with onset p- and coda -p is 3, and the number of stems with onset p- and coda -m is 11. Following previous researches (Frisch et al. 2004, Coetzee and Pater 2006), segments are categorized to major co-occurrence classes as follows: Labials {p, p^h, p', m, P}; Coronal Obstruents {c, c^h, c', t, t^h, t', s, s'}; Coronal Sonorants {l, n}; Dorsals {k, k^h, k', ŋ}. Coronal lenis codas -T, -S are treated as an intermediate category of Coronal Obstruents and Coronal Sonorants, since it is difficult to determine whether they belong to obstruent or sonorant. Onset-zero, coda-zero, and h are separated from these classes: onset-zero and coda-zero do not have strong co-occurrence restrictions except for combining with coda -h; as to h, see footnote 6.

Table 10. Onset-coda co-occurrences (O: onset, C: coda)

O \ C	p	p^h	P	m	t	t^h	c	c^h	s	s’	T	S	l	n	k	k^h	k’	ŋ	Ø	h	Totals
p	3		1	11	4	8	3	1	15		1	2	24	4	12	1	2	5	22		119
p^h				6		1							8					1	16		32
p’	1			6	1				1				5		1			2	7	1	25
m			2	4	3	4	5	3	8		1	1	25	3	15		1	5	37		117
t		1	2	9	4		1	3	6		3		22	3	7		3	11	18	1	94
t^h	2	2		1					2				5		3			9	23		47
t’				11	1				1				8	1	4			5	44	1	76
c	7	2	1	7		2	9	1	6			6	14	4	11			8	33	1	112
c^h			1	9			1						2	1	2			5	44		65
c’				4			2	1					1		11				21	1	41
s	3	6	1	21		3		1	3		1		29	9	9		2	2	31		121
s’	2			6	1				1				7		4				27	1	49
n	3	2	4	4	1	1	3	3	4		1	2	17	5	2		1	3	25	4	85
k	9	2	5	23	5	3	5		13		5	2	38	3	1		1	4	27		146
k^h				1									2	1				2	16		22
k’	1			5		2	1	2				1	8	2	1		2	4	28		57
Ø	4	4		11	1	3	1	2	8	1		2	24	4	9		1	4	24		103
h				3		2							4	3	2			3	18		35
Totals	35	19	17	142	21	29	31	17	68	1	12	16	243	43	94	1	13	73	461	10	1346

Table 11 shows the Observed and O/E values in the combination of each place of articulation, excluding Ø, h, and coronal lenis codas -T, -S. A value of O/E greater than 1 suggests no co-occurrence restriction, whereas a value of O/E less than 1 suggests the existence of co-occurrence restrictions. In some languages, pairs of identical consonants are unrestricted in co-occurrence, but Korean does not originally have many such pairs (p-p: 3, m-m: 4, t-t: 4, c-c: 9, s-s: 3, n-n: 5, k-k: 1, k’-k’: 2), and hence it is hard to say whether the exceptional behavior of identical consonant pairs exists in Korean or not. For this reason, we do not exclude these pairs from Table 11.

Table 11. Co-occurrence restrictions (O: onset, C: coda)

O \ C	Labial	Cor obs	Cor son	Dorsal	Totals
Labial	**34/52.0 (0.65 [0.74])**	57/40.6 (1.41)	69/68.3 (1.01)	45/44.1 (1.02)	205
Cor obs	98/88.5 (1.11)	**49/69.1 (0.71 [0.78])**	**106/116.3 (0.91 [0.95])**	96/75.1 (1.28)	349
Cor son	13/13.4 (0.97 [1.06])	12/10.5 (1.14)	22/17.7 (1.25)	**6/11.4 (0.53 [0.74])**	53
Dorsal	46/37.0 (1.24)	31/28.9 (1.07)	54/48.7 (1.11)	**15/31.4 (0.48 [0.60])**	146
Totals	191	149	251	162	753

The distribution is significant ($\chi^2 = 41.84$, $p < 0.001$). As a whole, co-occurrence of homorganic consonant pairs tends to be underrepresented except for Cor son–Cor son pairs. Unlike in many other languages, coronal obstruent and coronal sonorant seem not to constitute a separate class clearly in Korean. The pairs of Coronal sonorant onsets and Dorsal codas are also underrepresented (0.53 [0.74]), but this is probably an accidental gap, resulting from the fact that the only coronal sonorant onset is n- in Korean. In fact, the pairs of opposite order (Dorsal onset and Coronal sonorant coda) are overrepresented (1.11). Thus Korean has an OCP-Place constraint in monosyllabic stems, although it is not as strong as Arabic verbal roots where the co-occurrence of adjacent homorganic consonant pairs is extremely underrepresented.

As to complex codas, there are only 73 words with them, out of which 19 stems have onset Ø and 5 stems have onset h-. Hence it is difficult to draw any conclusion about an OCP-Place constraint based on this small data. However, the inventory of complex codas itself shows an OCP-Place constraint: in Korean complex codas -ks, -lh, -lk, -lm, -lp, -lp^h, -lP, -ls, -lt^h, -nc, -nh, -ps, coronals freely combine with any place of articulation, whereas labials and dorsals can combine only with coronals (s and l); complex codas such as *-mp and *-ŋk do not exist. This suggests the existence of an OCP-Place constraint in Korean as well as the unmarkedness of coronals: coronals often show asymmetric phonological phenomena as compared to other consonants in many languages (Kean 1975, the articles in Paradis and Prunet 1991, Davis 1992, and others).

3.2 Complexity restriction

Korean monosyllabic stems have a strong restriction in their complexity. Table 12 shows the distribution of stem structures, where numbers indicate the attested numbers of monosyllabic stems with each stem structure. (In Table 12, C = lax consonants, fricatives, nasals, a liquid, C^h = aspirated consonants, C' = tense consonants, CC, CC^h = complex codas, V = vowels, G = glides, V/GV = variations between ö and we. Lenis codas are excluded.) For example, there are 12 stems whose structure is V, and 484 stems whose structure is CVC.

Table 12. Stem structure

Onset \ Rhyme	V	VC	VC^h	VC’	VCC	VCC^h
Ø	12	50	5	1	16	1
C	177	484	45	9	40	3
C^h	89	62	3		1	
C’	116	106	5	2	9	
Totals	394	702	58	12	66	4

Onset \ Rhyme	GV	GVC	GVC^h	GVC’	GVCC	V/GV	VC/GVC
Ø	8	16	4	1	2	4	
C	8	32	4	1		26	2
C^h	9	1				1	
C’	2	6				9	1
Totals	27	55	8	2	2	40	3

(4) Examples

V: i- ‘roof (a house),’ i ‘tooth’ **VC**: ip- ‘wear,’ ip ‘mouth’ **VC^h**: əp^h- ‘turn upside down,’ ap^h ‘front’ **VC’**: is’- ‘exist, have’ **VCC**: anc- ‘sit,’ ols ‘reward’ **VCC^h**: ɨlp^h- ‘recite’ **GV**: je- ‘go (old style),’ jo ‘mattress’ **GVC**: jəl- ‘open,’ jəl ‘ten’ **GVC^h**: jat^h- ‘be shallow’ jəp^h ‘side’ **GVC’**: jək’- ‘knit’ **GVCC**: jalp- ‘be thin,’ jəlp- ‘be thin’ **V/GV**: ö-/we- ‘avoid,’ ö/we ‘cucumber’ **CV**: ka- ‘go,’ ka ‘edge’ **CVC**: mal- ‘stop,’ mal ‘horse’ **CVC^h**: mat^h- ‘smell,’ pat^h ‘field’ **CVC’**: tak’- ‘polish,’ pak’ ‘outside’ **CVCC**: malk- ‘be clear,’ hɨlk ‘soil’ **CVCC^h**: halt^h- ‘lick,’ holt^h- ‘scrape off’ **CGV**: pjə ‘rice plant,’ hwɛ ‘torch’ **CGVC**: kwal- ‘(flames) be furious,’ pjəl ‘star’ **CGVC^h**: pjət^h ‘sunlight,’ kjət^h ‘side’ **CGVC’**: kjək’- ‘experience’ **CV/CGV**: tö-/twe- ‘become,’ mö/mwe ‘grave’ **CVC/CGVC**: köm/kwem ‘favor,’ köm/kwem ‘support’ **C^{h}V**: c^{h}a- ‘kick,’ p^{h}a ‘scallion’ **C^{h}VC**: p^{h}al- ‘sell,’ k^{h}al ‘knife’ **C^{h}VC^h**: t^{h}op^h- ‘search for a way in a steep slope,’ p^{h}at^h ‘red bean’ **C^{h}VCC**: c^{h}ilk ‘arrowroot’ **C^{h}GV**: p^{h}jə- ‘spread,’ k^{h}jə ‘layer’ **C^{h}GVC**: p^{h}jəm ‘rice cake (jargon by ginseng collectors)’ **C^{h}V/C^{h}GV**: c^{h}ö/c^{h}we ‘clasp’ **C’V**: c’a- ‘be salty,’ t’ɛ ‘dirt’ **C’VC**: p’al- ‘wash,’ s’al ‘rice’ **C’VC^h**: c’oc^h- ‘chase,’ k’oc^h ‘flower’ **C’VC’**: k’ak’- ‘shave,’ k’ək’- ‘break’ **C’VCC**: c’alp- ‘be short,’ k’ɨlh- ‘boil’ **C’GV**: k’we- ‘thread,’ p’jə ‘bone’ **C’GVC**: p’jəm- ‘measure by fingers,’ p’jam ‘cheek’ **C’V/C’GV**: c’ö-/c’we- ‘expose,’ k’ö/k’we ‘wisdom’ **C’VC/C’GVC**: k’öm/k’wem ‘temptation’

Co-occurrence restriction in laryngeal features is observed: tense and aspirated consonants may not co-occur in one stem. Similar co-occurrence restrictions are observed in other languages (MacEachern 1997). In our data, there are only 10 exceptions (k’ak’- ‘shave,’ k’ək’- ‘break,’ c’oc^h- ‘chase,’ k’oc^h ‘flower,’ k’oc^h ‘fire (jargon)’ k’ɨt^h ‘edge,’ k’ɨt^h ‘a fold of cloth (‘incorrect form’),’ p^{h}at^h ‘red bean,’ t^{h}op^h- ‘search for a way in a steep slope,’ t^{h}op^h- ‘draw a skin of hemp with a scissors-like tool t^{h}op’), among which at least 4 words are attested as having a plain onset in MK (k’ak’- < kask-, k’ək’- < kəsk-, k’oc^h < koc, k’ɨt^h < kɨt^h).

The co-occurrence restriction in laryngeal features becomes more stringent when a coda is -CC^h or when a stem has a glide: no example is attested in these cases (*C^{h}VCC^h, *C^{h}GVC^h, *C^{h}GVC’, *C’VCC^h, *C’GVC^h, *C’GVC’). This fact suggests the existence of an accumulated complexity restriction (Anttila 2007). Assuming C’, C^h, diphthongs and complex codas as a complex segment uniformly, more than two complex segments cannot co-occur in Korean monosyllabic stems. For example, C^{h}GVC^h is unavailable because it violates *COMPLEX three times (onset C^h, GV, and coda C^h). Similarly, C’VCC^h is prohibited because it contains three violations (C’, complex coda, and C^h). The fact that there is no rhyme -GVCC^h is also due to the ban against three violations. (Lenis codas -P, -T, -S may belong to a complex segment as well, because they rarely combine with C^h and C’: out of 45 lenis codas, only 2 (c^{h}uP- ‘be cold,’ k’ɨS- ‘pull’) combine with C^h/C’.)

This is basically the same idea as a self-conjoined constraint No-φ^2 (Itô and Mester 2003), which prohibits the co-occurrence of the structure φ with itself in domain δ. The different point in Korean case is that three, not two, violations are prohibited (*COMPLEX^3_m, m = morpheme), and that the effect of the self-conjoined constraint is gradient. That is, two violations as seen in **C’**V**C^h**, C**GVC^h**, **C^h**V**CC**, etc. are not completely disallowed in Korean, although these examples are relatively rare. This means that the ban against two violations of *COMPLEX is weaker than against three violations and hence such structures could narrowly occur.

Thus co-occurrence restrictions in Korean stem structures can be uniformly explained by a gradient constraint *COMPLEX, which treats different classes such as C’ C^h, GV, and CC as one complex segment.

4. Predictable distribution of aspirated consonants

In Korean monosyllabic stems, aspirated onsets mainly combine with coda Ø or sonorant codas, whereas aspirated codas tend to combine with obstruent onsets. Table 13 shows these correlations (O = obstruent, R = sonorant, L = lenis coda, lO/lC^h = complex codas (l + obstruent/C^h)).

Table 13. The distribution of an aspirated consonant in a stem

Onset	C^h	C^h	C^h	C^h	C^h	O	R	Ø	O	Ø	C^h
Coda	Ø	R	O	L	lO	C^h	C^h	C^h	lC^h	lC^h	C^h
Totals	99	53	10	1	1	41	13	9	3	1	3

In reality, the position of an aspirated consonant in a stem is highly predictable in Korean.[4] The structure of Korean monosyllabic native stems can be stated simply[5]: there is only one aspiration per stem as a rule, and it appears on the rightmost obstruent (= coda); if the rightmost segment is either Ø, sonorants (l, m, n, ŋ), lenis codas (-P, -T, -S), or fricatives (s, h), which cannot realize aspiration by itself, then the [+asp] is moved to another position (= onset). The constraint against assigning aspiration in sonorants/fricatives/lenis codas is represented as $*R^h$ here. The constraint ranking is $*C^hVC^h > *R^h >$ Max-asp > Align-asp-R. (5) and (6) are the examples. These tableaux describe the output forms of the stem-level stage in the sense of OT-based lexical phonology model (Kiparsky 2000).

(5) pat^h 'field'

Input: $/p^hat^h/$	$*C^hVC^h$	$*R^h$	Max-asp	Align-asp-R
a. p^hat^h	* !			*
☞ b. pat^h			*	
c. p^hat			*	* !
d. pat			** !	

(6) c^him 'saliva'

Input: $/c^him^h/$	$*C^hVC^h$	$*R^h$	Max-asp	Align-asp-R
a. c^him^h	* !	*		*
b. cim^h		* !	*	
☞ c. c^him			*	*
d. cim			** !	

Furthermore, there is a hierarchy among obstruents p, c, t, k, in bearing aspiration. As mentioned above, k^h is quite rare in Korean monosyllabic stems, and in fact, the avoidance of k^h is observed in the distribution of aspiration. We assume a constraint $*k^h$ here: k is not the target of $*R^h$, since it can bear an aspiration when another segment in a stem is R. The revised constraint ranking is $*C^hVC^h > *R^h >$ Max-asp $> *k^h >$ Align-asp-R. (7) -

[4] As mentioned above, the relative segmental frequencies of lax consonants p-, t-, c-, k- are the mirror image of aspirated consonants p^h-, t^h-, c^h-, k^h-. Aside from the input forms in the tableaux of this section, this fact and the predictability of aspiration may suggest that aspirated consonants originally derived from lax consonants by some sound change in Korean history.

[5] Itô and Mester (1989) provides similar distributional generalizations concerning mimetic palatalization in Japanese.

(9) show the examples.

(7) t^hək 'chin'

Input: /t^hək^h/	*C^hVC^h	*R^h	Max-asp	*k^h	Align-asp-R
a. t^hək^h	* !			*	*
b. tək^h			*	* !	
☞ c. t^hək			*		*
d. tək			** !		

(8) kat^h- 'be identical'

Input: /k^hat^h-/	*C^hVC^h	*R^h	Max-asp	*k^h	Align-asp-R
a. k^hat^h-	* !				*
☞ b. kat^h-			*		
c. k^hat-			*	* !	*
d. kat-			** !		

(9) k^hoŋ 'bean'

Input: /k^ho$ŋ^h$/	*C^hVC^h	*R^h	Max-asp	*k^h	Align-asp-R
a. k^ho$ŋ^h$	* !	*		*	*
b. ko$ŋ^h$		* !	*		
☞ c. k^hoŋ			*	*	*
d. koŋ			** !		

The same is true for complex codas as in (10) and (11).

(10) halt^h- 'lick'

Input: /h^halt^h-/	*C^hVC^h	*R^h	Max-asp	*k^h	Align-asp-R
a. h^halt^h-	* !	*			*
☞ b. halt^h-			*		
c. hal^ht-		* !	*		*
d. h^halt-		* !	*		*
e. halt-			** !		

(11) c^hilk 'arrowroot'

Input: /c^hilk^h/	*C^hVC^h	*R^h	Max-asp	*k^h	Align-asp-R
a. c^hilk^h	* !			*	*
b. cilk^h			*	* !	
c. cil^hk		* !	*		*
☞ d. c^hilk			*		*
e. cilk			** !		

(12) Other examples
ip^h ‘leaf,’ ɨlp^h- ‘recite,’ coc^h- ‘follow,’ cip^h ‘straw,’ c^{h}əl ‘season,’ c^{h}uk- ‘be damp,’ c^{h}uP- ‘be cold,’ kap^h- ‘pay back,’ kjət^h ‘side,’ k^{h}al ‘knife,’ k^{h}i ‘height,’ k^{h}ɨ- ‘be big,’ mit^h ‘bottom,’ nac^h ‘face,’ nop^h- ‘be high,’ pɛt^h- ‘spit,’ pjət^h ‘sunlight,’ pic^h ‘light,’ p^{h}a ‘scallion,’ p^{h}i ‘blood,’ p^{h}jə- ‘spread,’ p^{h}ul ‘grass,’ p^{h}ul ‘paste,’ səp^h ‘gusset,’ suc^h ‘charcoal,’ təc^h ‘trap,’ təp^h- ‘cover,’ t^{h}a- ‘ride’, t^{h}as ‘reason,’ t^{h}ə ‘base,’ t^{h}əl- ‘dust, hit,’ etc.

Aside from 3 C^{h}VC^h stems, the exceptions are only c^{h}ac- ‘look for’ and t^{h}op ‘saw.’ [6]

Thus in the stem-level stage, the avoidance of k^h in Korean is extreme. Is the dispreference for k^h typologically common? In the next section, we investigate the typological tendency of voiceless aspirated obstruents.

5. Typology of aspiration

The subject of investigation is 64 languages from the database in Maddieson 1984. Our data is based on the languages where voiceless aspirated obstruents contrast with plain voiceless obstruents at least in one place of articulation. Languages which have aspiration only phonetically are excluded. Voiceless sonorants, which are also characterized by the feature [spread glottis], are not examined here. As to the correlation between voiceless sonorants and aspirates, see Halle and Stevens 1971, Itô and Mester 1989, Lombardi 1994, Clements 1985, 2003. (In Table 14, P = labial stops, T = coronal stops, K = velar stops, Q = uvular stops, Ts = coronal affricates, Kx = velar affricates, Pf = labio-dental affricates, S = coronal fricatives. Coronal includes dental, alveolar, retroflex, and palatal.)

[6] Davis and Cho (2003) provides an analysis of the distribution of h and aspirated stops in Korean, showing that their word-level distribution is parallel to an extent in that neither h nor aspirated stops appear in coda position while both can appear in word-initial position. In fact, similar parallelism between h and aspirated obstruents are observed in the stem-level phonotactics. a) h rarely combines with aspirated obstruents or h. (6 out of 64 stems show the co-occurrence of h and h/aspirated obstruents: hot^h ‘single thing,’ hɨt^h- ‘scatter,’ halt^h- ‘lick,’ holt^h- ‘scrape off,’ hult^h- ‘draw through one’s hands,’ hɨnh- ‘be common.’ The aspirated obstruent with which h combines is only t^h.) b) Onset h- basically combines with coda Ø, sonorants or -k, and does not combine with lax obstruents other than -k, whereas coda containing h combines with lax obstruent onsets k-, t-, c- as well as other onsets. E.g.) hjə ‘tongue,’ hwal ‘bow,’ hok ‘bump,’ hɨlk ‘soil,’ kolh- ‘rot,’ tah- ‘touch,’ coh- ‘be good,’ ilh- ‘lose,’ manh- ‘be numerous.’ Thus the distribution of h and aspirated obstruents is parallel, but h is not completely parallel to aspirated obstruents in that h often combines with tense consonants. E.g.) k’ɨlh- ‘boil,’ s’ah- ‘pile up,’ t’ulh- ‘excavate,’ c’ih- ‘pound.’

Table 14. Frequency of languages with voiceless aspirated obstruents

Unaspirated	Aspirated	No	Languages
P, T, K, Ts, S	P^h, T^h, K^h, Ts^h	28	Amoy, Bengali, Changchow, Cofan, Dakota, E. Armenian, Fuchow, Georgian, Hakka, Jingpho, Kan, Kashmiri, Kharia, Khmer, **Korean**, Lakkia, Lungchow, Mandarin, Mundari, Punjabi, Swahili, Taishan, Tamang, Tarascan, Thai, Wiyot, Yao, Yuchi
P, T, K, Q, Ts, S	P^h, T^h, K^h, Q^h, Ts^h	7	Burushaski, Haida, Jaqaru, Lahu, Lak, Quechua, Sui
P, T, K, S	P^h, T^h, K^h	6	Cham, Daribi, Khasi, Nambakaengo, S. Nambiquara, Yay
P, T, K, Ts, S	P^h, T^h, K^h	5	Azerbaijani, Igbo, Otomi, Saek, Sedang
P, T, K, Ts, S	P^h, T^h, K^h, Ts^h, S^h	2	Burmese, Mazahua
P, T, K, Ts, S	P^h, T^h	2	Tiddim Chin, Tiwa
P, T, K, Q, Ts, S	P^h, T^h, K^h, Ts^h	1	Hindi-Urdu
P, T, K, Q, Ts, S	P^h, T^h	1	Wintu
P, T, K, Q, S	P^h, T^h, K^h, Q^h, Ts^h	1	Gilyak
P, T, K, Ts, S	T^h, K^h, Ts^h	1	Navaho
P, T, K, Ts, S	T^h, Ts^h	1	Tolowa
P, T, K, Ts, S	Ts^h, Kx^h	1	Chipewyan
P, T, K, S	P^h, T^h, K^h, S^h	1	Karen
P, T, K	P^h, T^h, K^h	1	Kunjen
P, T, Q, Ts, S	P^h, T^h, Q^h	1	Kirghiz
P, T, Ts, S, Pf	P^h, T^h, K^h, Ts^h, Pf^h	1	Beembe
P, T, S	K^h, Ts^h	1	Zuni
T, K, Ts, S	T^h, K^h, Ts^h	1	Wichita
T, K, S	T^h	1	Vietnamese
T, Q, Ts, S	T^h, Ts^h	1	Hupa

The most frequent type (28 languages) is "P, T, K, Ts, S vs. P^h, T^h, K^h, Ts^h." That is, stops and affricates have aspirated/unaspirated oppositions whereas only S does not. The second most frequent type (7 languages) is similar to this type, adding the uvular opposition. The third most frequent type (6 languages) is a subset of the most frequent type in that this case just lacks an affricate, and only coronal fricative S belongs to the unaspirated series. In 5 languages such as Azerbaijani, Igbo, Otomi, etc., stops have aspirated/unaspirated oppositions whereas neither affricates nor fricatives have. Thus the aspirated series is more restricted than the unaspirated series in both place of articulation and articulatory manner. (13) shows these hierarchies, along with combined consequence in (13c). Parentheses indicate the total numbers of languages. Labio-dental Pf is included into Labial.

(13) Hierarchies of the aspirated series

a. Coronal (64) > Labial, Velar (57) > Uvular (9)

b. Stop (63) > Affricate (46) > Fricative (3)

c. T^h (62) > P^h (57) > K^h (56) > Ts^h (46) > Q^h (9) > S^h (3) > Pf^h, Kx^h (1)

In plain voiceless plosives (pulmonic egressive stops), the presence of /p/ implies the strong likelihood of the presence of /k/, which similarly implies the presence of /t/ (dentals and alveolars) (Maddieson 1984). Although the frequencies of labials and velars in (13a) are the same, the hierarchy in the aspirated series is similar to plain voiceless plosives, and different from ejectives. The fact that aspirated affricates are coronals as a rule also coincides with the cross-linguistic tendency in plain voiceless obstruents.

Many languages have the aspiration opposition in both stops and affricates, excluding only fricatives. This is different from the hierarchy in plain voiceless obstruents, where affricate implies fricative (*affricate > *fricative). This means that a closure phase is required in producing aspiration, and that the frication phase of affricate does not significantly result in cross-linguistic disfavor of aspirated affricate. Fricative does not have a closure phase at all, and hence it is strikingly underrepresented in the aspirated series.

Still, a question is raised: if the unaspirated/aspirated distinction just depends on the difference of voice onset timing (VOT), S vs. S^h should be as possible as Ts vs. Ts^h theoretically. Other than closure, what are the crucial acoustic factors which make Ts vs. Ts^h more distinctive than S vs. S^h? On the other hand, how is an aspiration realized in affricate and fricative? Is there a clear distinction between aspiration and frication in aspirated affricate/fricative? Accounting for these problems with acoustic data is a task for future research.

Typologically, K^h is not disfavored as seen in (13c). Given this, the extreme underrepresentation of k^h in Korean is unusual. However, this is probably due to the acoustic similarity between [k^h]/[q^h] and fricatives [x]/[χ], which may have resulted in the sound change from [k^h] to [h] in Korean most stems historically, since Korean doe not have [x]/[χ]. In fact, these fricatives always exist in the languages which have k/q but do not have k^h/q^h.

Table 15. Languages without k^h/q^h

Language	Unaspirated	Aspirated	Fricative (unaspirated)
Tiddim Chin	p, t, ts, k	p^h, t^h	s, x, h
Tiwa	p, “t,” tʃ, k, k^w	p^h, “t^h”	“s,” x, x^w, h
Tolowa	p, “t,” tʃ, k, k^w	“t^h,” tʃ^h	s, ʃ, ṣ, x, x^w, h
Vietnamese [7]	t, c, k	t^h	f, s, x, h
Wintu	p, t, tʃ, k, q	p^h, t^h	f, θ, s, x, χ, h
Hupa	t, ts, tʃ, c, q	t^h, ts^h, tʃ^{wh}, c^h	s, x, x^w, h, h^w

Thus the extreme underrepresentation of k^h in Korean can be typologically unnatural, if that resulted from the confusion with velar/uvular fricatives.

6. Conclusion

In this paper, we have argued for the segmental frequency, co-occurrence restrictions and the distribution of aspiration in Korean monosyllabic stems, along with a typological survey of aspiration.

The segmental frequency in Korean monosyllabic stems coincides with a cross-linguistic markedness hierarchy: unmarked segments/classes are more attested than marked ones, such as the lax series > the aspirated/tense series, simplex codas > complex/lenis codas, monophthongs > diphthongs, peripheral vowels > central vowels. Our data also shows some general tendencies: tense consonants (voiceless laryngealized obstruents) have similar hierarchic patterns to ejectives; the structure of monosyllabic stems correlates with the Sonority Cycle; verbs have more complex coda inventory than nouns, which is probably due to the Base-Identity effect.

Two co-occurrence restrictions are observed in Korean. First, as in many languages which are genetically unrelated, Korean has an OCP-Place constraint. Second, Korean has a gradient complexity constraint that restricts stems composed of more than one complex segment, such as aspirated/tense consonants, diphthongs, complex codas.

In Korean monosyllabic stems, the position of aspiration is predictable. It occupies the coda position as a rule, but moves to the onset position when the coda is either a segment which cannot bear aspiration or -k. This distribution is explained by the constraint ranking $*C^hVC^h > *R^h >$ Max-asp > $*k^h$ > Align-asp-R.

k^h is highly underrepresented in the Korean lexicon. To place this extreme underrepresentation in perspective, we conducted a typological survey of language frequency for aspiration contrasts. The implicational hierarchy is Coronal > Labial, Velar > Uvular for place and Stop > Affricate > Fricative for manner. The extreme underrepresentation of k^h in Korean is not due to

[7] Historical studies have revealed that Vietnamese p^h and k^h changed to f, x respectively.

the cross-linguistic dispreference for k^h but probably results from potential confusion with a velar/uvular fricative.

REFERENCES

ALBRIGHT, ADAM. 2002a. The identification of bases in morphological paradigms. Ph.D.dissertaion, UCLA.

———. 2002b. The lexical bases of morphological well-formedness. Sabrina Bendjaballah, Wolfgang U. Dressler, Oskar E. Pfeiffer and Maria D. Voeikova (Eds.) *Morphology 2000*. John Benjamins.

———. in press. Explaining universal tendencies and language particulars in analogical change. Jeff Good (Ed.) *Language Universals and Language Change*. Oxford University Press.

ALBRIGHT, ADAM and BRUCE HAYES. 2002. Modeling English Past Tense Intuitions with Minimal Generalization. Maxwell, Michael (Ed.) *Proceedings of the Sixth Meeting of the ACL Special Interest Group in Computational Phonology. Philadelphia*, July 2002. ACL.

ANTTILA, ARTO. 2007. Gradient phonotactics in Optimality Theory. Unpublished Stanford University ms.

CHO, TAEHONG, SUN-AH JUN and PETER LADEFOGED. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics 30*, 193-228.

CHOMSKY, NOAM and MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper & Row.

CLEMENTS, GEORGE N. 1985. The geometry of phonological features. *Phonology Yearbook 2*, 225-252.

———. 1990. The role of the sonority cycle in core syllabification. In Kingston, John and Mary E. Beckman (Eds.). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press.

———. 2003. Feature economy in sound systems. *Phonology 20*, 287-333.

COETZEE, ANDRIES W. and JOE PATER. 2006. Lexically ranked OCP-Place Constraints in Muna. Ms., University of Michigan and University of Massachusetts, Amherst.

DAVIS, STUART. 1992. Investigating English phonotactic constraints using a computerized lexicon. In Ursula Klenk (Ed.). *Computatio Linguae*, 1-15. Stuttgart: Franz Steiner Verlag.

DAVIS, STUART and MI-HUI CHO. 2003. The distribution of aspirated stops and /h/ in American English and Korean: an alignment approach with typological implications. *Linguistics 41-4*, 607-652.

FRISCH, STEFAN A., JANET B. PIERREHUMBERT and MICHAEL B. BROE. 2004. Similarity Avoidance and the OCP. *Natural Language and Linguistic Theory 22*, 179-228.

GREENBERG, JOSEPH H. 1950. The Patterning of Root Morphemes in

Semitic. *Word 6*, 162-181.
———. 1970. Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics 36-2*, 123-145.
HALLE, MORRIS and KENNETH N. STEVENS. 1971. A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Progress Report 101*, 198-212.
ITÔ, JUNKO and ARMIN MESTER. 1989. Feature predictability and underspecification: palatal prosody in Japanese mimetics. *Language 65-2*, 258-293.
———. 2003. *Japanese Morphophonemics: Markedness and Word Structure*. Cambridge: The MIT Press.
IVERSON, GREGORY K. 1983. Korean s. *Journal of Phonetics 11*, 191-200.
JUN, SUN-AH. 1993. The phonetics and phonology of Korean prosody. Ph.D. dissertation, Ohio State University.
KAGAYA, RYOHEI. 1974. A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics 2*: 161-180.
KAWAHARA, SHIGETO, HAJIME ONO and KIYOSHI SUDO. 2006. Consonant Cooccurrence Restrictions in Yamato Japanese. *Japanese/Korean Linguistics 14*, 27-38.
KEAN, MARIE-LOUISE. 1975. The Theory of Markedness in Generative Grammar. Ph.D. dissertation, MIT.
KENSTOWICZ, MICHAEL. 1996. Base-Identity and Uniform Exponence: Alternatives to Cyclicity. In Durand, Jacques and Bernard Laks (Eds.). *Current Trends in Phonology: Models and Methods*, 363-393. Stanford: University of Stanford.
KIPARSKY, PAUL. 2000. Opacity and cyclicity. *The Linguistic Review 17*, 351-366.
LADEFOGED, PETER and IAN MADDIESON. 1996. *The sounds of the world's languages*. Cambridge: Cambridge University Press.
LEE, IKSOP and ROBERT RAMSEY. 2000. *The Korean Language*. New York: State University of New York Press.
LOMBARDI, LINDA. 1994. *Laryngeal features and laryngeal neutralization*. New York: Garland.
MACEACHERN, MARGARET R. 1997. Laryngeal Cooccurrence Restrictions. Ph.D. dissertation, UCLA.
MADDIESON, IAN. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
MCCARTHY, JOHN J. 1986. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry 17*, 207–263.
———. 1988. Feature Geometry and Dependency: A Review. *Phonetica 45*, 84-108.
———. 1994. The phonetics and phonology of Semitic pharyngeals. In Keating, Patricia A. (Ed.). *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology 3*, 191-233. Cambridge: Cambridge

University Press.
MESTER, ARMIN. 1986. Studies in Tier Structure. Ph.D. dissertation, University of Massachusetts, Amherst.
MIKHEEV, ANDREI. 1997. Automatic rule induction for unknown-word guessing. *Computational Linguistics 23*, 405-423.
PADGETT, JAYE. 1995. *Stricture in Feature Geometry*. Stanford: Center for the Study of Language and Information.
PARADIS, CAROLE and JEAN-FRANÇOIS PRUNET. 1991. *Phonetics and Phonology 2: The Special Status of Coronals: Internal and External Evidence*. San Diego: Academic Press.
PARK, HANSANG. 1999. The phonetic nature of the phonological contrast between the lenis and fortis fricatives in Korean. *Proceedings of the 14th International congress of Phonetic science (ICPhS99) 1*, 424-427.
PIERREHUMBERT, JANET B. 1994. Syllable Structure and Word Structure. *Papers in Laboratory Phonology III*, 168-188. Cambridge: Cambridge University Press.
PRINCE, ALAN and PAUL SMOLENSKY. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University and University of Colorado, Boulder.
VAUX, BERT and BRIDGET SAMUELS. 2005. Laryngeal markedness and aspiration. *Phonology 22*, 395-436.

Chiyuki Ito
Research Institute for Languages and Cultures of Asia and Africa
Tokyo University of Foreign Studies
3-11-1 Asahi-cho, Fuchu-shi, Tokyo 183-8534, Japan
e-mail: chiyukit@aa.tufs.ac.jp