Neutralizing Aplosivation and anti-homophony in Korean

Daniel Silverman
SJSU
daniel.silverman@sjsu.edu
UCB Phorum 12:00-1:00pm Monday, September 24, 2007

> *"…[T]here is no doubt that in some way or other, linguistic systems respond to change in ways that maintain meaning—more or less"* -William Labov (1994)

- Obstruent coda laryngeal neutralization in Korean is a case *par excellence* of neutralization, and suggests that phonological systems are not influenced by functional considerations such as *contrast maintenance*.
- However, a language could *never* evolve towards a state in which its communicative function becomes genuinely eroded.
- The real question: **does neutralization induce extensive homophony?** (Much like Martinet's "functional load".)
- Even more pointedly: **does homophony induce semantic ambiguity?**
- Grammatical and real-world knowledge both play great roles in disambiguating homophonous forms, but in many cases neutralization is tolerated *exactly because* the phonology and/or morphology provide(s) an "escape hatch"; the threat of ambiguity is curtailed.
- The proposal: in Korean, coda aplosivization developed *exactly because* the language acquired morphological means to countervail the threat of homophony/ambiguity.
- Neutralization in Korean is not counter-functional: language does not evolve towards counter-functional ends.

## Structure of this presentation
- Exemplify the aplosivization/neutralization pattern of Korean coda obstruents.
- Investigate the extent of potential homophony.
- Trace the historical development of the pattern.
- Consider phonetic and morphological motivation for the sound change.

(1) Obstruent/h prevocalic and non-prevocalic contrastive values in Korean (shaded values are not contrastive in root-final position):

Lexically prevocalic					Lexically non-prevocalic
Plain	Voiced	Aspirated	Tensed		Aplosive
p	b	p^h	p'		$p^˺$
t	d	t^h	t'	alternates with	
tʃ	dʒ	$tʃ^h$	tʃ'		$t^˺$
s		h	s'		
k	g	k^h	k'		$k^˺$

(2) A few examples (adapted from Jun, in progress):

Root-Final C	Plosive		Aplosive	gloss
	(i) Locative (-e)	(ii) Nominative (-i)	(iii) Isolation Form	
Labial	pab̲-e ip̲ʰ-e	pab̲-i ip̲ʰ-i	pap̚ ip̚	rice leaf
Alveolar	os̲-e pat̲ʰ-e nad̲ʒ-e pit̲ʃʰ-e	oʃ-i patʃʰ-i nad̲ʒ-i pit̲ʃʰ-i	ot̚ pat̚ nat̚ pit̚	clothes field day light
Velar	kug̲-e pak̲'-e puək̲ʰ-e	kug̲-i pak̲'-i puək̲ʰ-i	kuk̚ pak̚ puək̚	soup outside kitchen

(3) Potential neutralizations among coronal obstruents in Korean (alternating values; "mutant dragonfly" display):
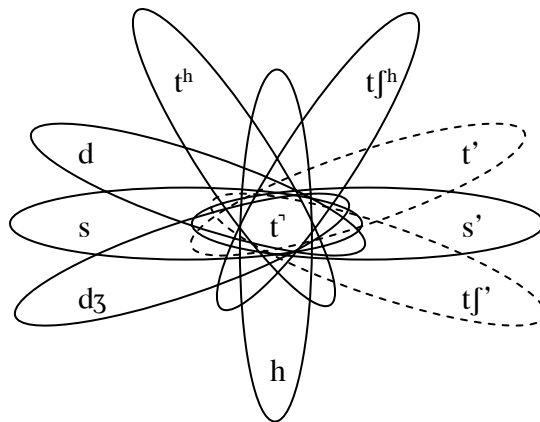
(4) Neutralization may induce homophony—a potentially counter-functional development:
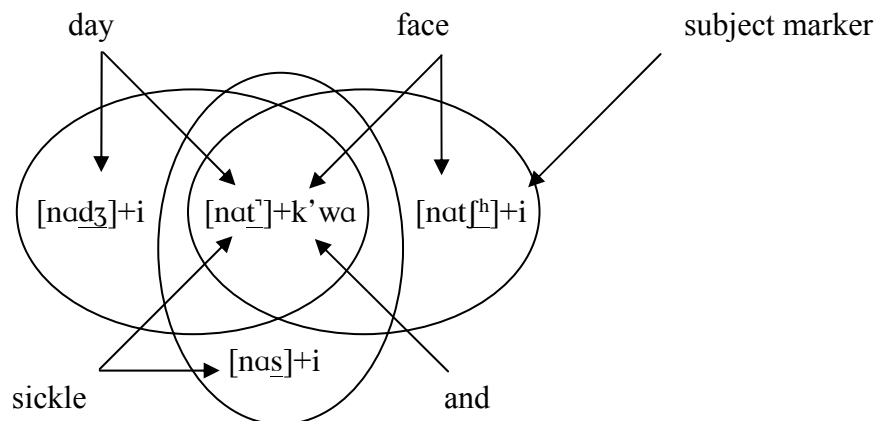
- Moreover, nouns often appear unsuffixed (case markers, for example, are variably present)
- Q1: Shouldn't this neutralization induce massive homophony?
- Q2: Are Korean speakers constantly misunderstanding each other?
- A1: No; A2 No. The potential for homophony is not manifested to any significant degree.

(5) Distribution of root-final obstruents for 35,908 nouns, from The Sejong Project, a 1.5 million word corpus (http://sejong.or.kr/english/index.html) (with invaluable assistance from Jongho Jun and Adam Albright)

Labial ([pˀ])		Alveolar ([tˀ])		Velar ([kˀ])	
Value	Number of words	Value	Number of words	Value	Number of words
p/b	1164	t/d	2	k/g	3508
pʰ	38	tʰ	63	kʰ	12
p'	0	t'	0	k'	4
		tʃ	10		
		tʃʰ	74		
		tʃ'	0		
		s	257		
		h	0		
		s'	0		
Totals:	1202		406		3524

- Clearly, the opportunities for neutralization are rather limited, and the chances of homophony are extremely low.

(6) List of potential homophonic noun sets in the corpus involving final obstruent/h neutralization (IPA to follow)

	IPA	Transliteration	Token Freq. (out of 1,353,602 noun tokens)	Hangul	Source	Gloss
1	tˀ]	{Ceoc}	44	젖	Native	breast/milk
		{Ceos}	5	젓	Native	Salted fish
2	pˀ]	{Cip}	9	집	Native	house
		{Ciph}	19	짚	Native	straw
3	tˀ]	{Mas}	392	맛	Native	taste
		{Math}	3	맡	Native	???
4	tˀ]	{Nac}	294	낮	Native	day
		{Nach}	91	낯	Native	face

		{Nas}	25	낫	Native	sickle
5	t˺]	{Pat}	2	받	Native	???
		{Path}	380	밭	Native	field
6	t˺]	{Pic}	128	빚	Native	debt
		{Pich}	726	빛	Native	light
		{Pis}	6	빗	Native	comb
7	t˺]	{Pyeos}	6	볏	Native	crest (of a fowl)
		{Pyeoth}	32	볕	Native	sunshine
8	t˺]	{Such}	14	숯	Native	charcoal
		{Suth}	7	숱	Native	(hair) density
9	t˺]	{Teoch}	31	덫	Native	trap
		{Teos}	4	덧	Native	a short time
10	p˺]	{ip}	1139	입	Native	mouth
		{iph}	158	잎	Native	leaf

- Ten (10) sets out of a total of 35,908 nouns in the corpus.
- Usually, low token frequency for one or both/all members.

- As for verb roots, they are obligatorily suffixed, and many of these suffixes are vowel-initial. Induced homophony is, consequently, extremely rare indeed.

(7) Distribution of root-final obstruents/h for 3801 verbs, from The Sejong Project.

Labial ([p˺])		Alveolar ([t˺])		Velar ([k˺])	
p/b	216	t/d	61	k/g	106
p^h	18	t^h	31	k^h	0
p'	0	t'	0	k'	12
		tʃ	72		
		$tʃ^h$	4		
		tʃ'	0		
		s	97		
		h	113		
		s'	9		
Totals:	234		387		118

(8) List of potential homophonic verb sets in the corpus involving final obstruent neutralization.

	IPA	Transliteration	Token Freq. (out of 644,403 verb tokens)	Hangul	Source	Gloss
1	tʼ]	{Cac}	82	잦		
		{Cas}	4	잣		
2	tʼ]	{Ceoc}	202	젖		
		{Ceos}	66	젓		
3	pʼ]	{Cip}	172	집		
		{Ciph}	125	짚		
4	tʼ]	{Cic}	20	짖		
		{Cis}	1143	짓		
		{Cith}	171	짙		
5	pʼ]	{eop}	73	업		
		{eoph}	32	엎		
		{eops}	14290	없		
6	tʼ]	{ic}	426	잊		
		{is}	548	잇		
		{iss}	47440	있		
7	tʼ]	{Kip}	15	깁		
		{Kiph}	587	깊		
8	tʼ]	{Kuc}	19	궂		
		{Kut}	167	굳		
9	kʼ]	{Muk}	112	묵		
		{Mukk}	155	묶		
10	tʼ]	{Nac}	300	낮		
		{Nas}	242	낫		
11	tʼ]	{Pic}	166	빛		
		{Pis}	15	빗		
12	kʼ]	{Sok}	49	속		
		{Sokk}	4	솎		
13		{Teop}	106	덥		

	p˺]	{Teoph}	160	덮	

- Thirteen (13) sets out of 3801 verbs in the corpus.
- Usually, low token frequency for one or both/all members.
- There are quite a few more potential verbs homophones, ratio-wise (13 sets/3801 words vs. 10 sets /35,908 words)

- Q: Isn't it remarkable that the system has so little homophony?
- A: Well, it's not magic…

Let's delve into history a bit…

(9)

1300 years ago	1100 years ago	500 years ago	400 years ago
Sino-Korean vocabulary, with its aplosive final stops, is well-established, supplanting many native nouns with two-root compounds.	Root-final [tʰ], [tʃ], [tʃʰ], and [s] are still found in final position, and before consonants, in native Korean	Influenced by the Sino-Korean root-structure, [tʰ], [tʃ], and [tʃʰ] become [t˺] in these positions. But verbs alternate, and Sino-Korean nouns are compounds.	[s] alternates with [t˺] in these positions as well. Still, there are few adverse functional consequences.

- Korean has witnessed a massive influx of Chinese words in its distant past, which served to supplant a significant portion of its native vocabulary, particularly its noun inventory.
- During the era of borrowing, Chinese only had [p˺, t˺, k˺, m, n, ŋ] in root-final position, while native Korean had all the contrasts that are now neutralized ([t, tʰ, s, tʃ, tʃʰ]); I assume that plain final "t" was unreleased: [t˺]. These root-final consonants may even have been present in word-final position, though their phonetic values are not known (e.g. Martin 1992).
- The influx of Chinese nouns into Korean eventually led to a reduced set of word-final obstruents, even for native words, which previously possessed releases.
- These changes rendered the native Korean vocabulary more similar in structure to the Sino-Korean vocabulary: final stops in the Sino-Korean vocabulary were always unreleased, and so the evolution towards aplosive final stops in the native vocabulary rendered the two systems more similar to each other in this respect.
- This is a rather common phenomenon: upon the influx of foreign elements, the phonological properties of a language may change; incoming patterns may eventually modify, or even supplant native ones.
- Why didn't this induce massive homophony?
- This reduction in contrastive sounds was offset by the compounding process, which greatly increased the opportunity for nouns to contrast with each other.
- Most Korean nouns of Chinese origin are actually compounds of two Chinese roots.

(10) Schematic: Native Korean nouns: Sino-Korean nouns:
        XYt$^{\textipa{?}}$ ABC+MNt$^{\textipa{?}}$
        XYt^{h} DEF+MNt$^{\textipa{?}}$
        XYs   GHI+MNt$^{\textipa{?}}$
        XYtʃ   JKL+MNt$^{\textipa{?}}$ …etc.

- For example, [ho], meaning "good", cannot stand on its own, but it often combines with other roots: [ho+gi] good opportunity, [ho+sa], happy event.
- So, any limitations on the number of possible contrasts imposed by the smaller inventory of Chinese root-final consonants was offset by these roots' combination and recombination into new and varied compounds.
- In its pre-history, Korean was probably perfectly happy with its many root-final consonants that always manifested themselves because they were released.
- However, due to the influx of Chinese, with its system of aplosive root-final stops, the possibility of the sound change in the native system was "encouraged", and ultimately actualized.
- That is, since the extensive Sino-Korean vocabulary had a fairly simple system of root-final consonants, this eventually had the effect of triggering a change in the native system of word-final consonants (which is a common sort of sound change even without any external "encouragement").
- So when native Korean roots stood alone, and when they were followed by a consonant, they eventually came to conform to the Sino-Korean pattern that was so extensive.
- It's likely that aplosive stops became so pervasive only because communication was not adversely affected, due to the compounding of Sino-Korean nouns. Release-loss was tolerated exactly because of the compounding process.
- Moreover, there were few sets that minimally contrasted in terms of coda laryngeal values anyway.
- The new prevalence of nouns with aplosive root-final stops may have been accompanied by the loss of releases in the verb system as well, which was tolerated because the many vowel-initial suffixes which accompany verbs produce alternations that salvage the contrasts that are dependent on the stop release.
- It is almost unimaginable that Korean would have tolerated the development of aplosive stops if it would have resulted in extensive homophony/ambiguity of word meaning.
- Q1: Is it just a fortuitous coincidence that both verbs and nouns—for completely different reasons—are so rarely homophonous?
- Q2: Could this situation have been brought about by design, in the sense that Chinese nouns were intentionally turned into compounds of two morphemes in order to increase the number of possible contrasts, and the replacement of native Korean verbs with single Chinese roots was intentionally resisted?
- A1: No; A2: No. There is no intention involved. Rather, the Korean phonological system simply evolved passively as a consequence of its communicative function (see Silverman 2006 for details).

- **As a consequence of SK compounding, the process under scrutiny is still underway today: the coda system in nouns is simplifying and regularizing, with very few counter-functional consequences: compounding has allowed the it to simplify (see Albright *ms.*, Ito *ms.*, Jun *ms.*, Kang *to appear*).**

Final remarks: What does the Korean pattern suggest?

Little:
- Despite the potential for neutralization in Korean, actually, there are minimal counter-functional consequences.

Medium:
- Neutralization is counter-functional only to the extent that it induces not only homophony, but semantic ambiguity; semantic ambiguity is thus argued to be a stealthy but powerful factor in sound change. ("From small things, big things one day come".)
- Neutralization is likely to never induce significant homophony/ambiguity of word meaning.
- Neutralization may become extensive, but only provided that significant homophony/ambiguity is not induced.

Big:
- Such patterns may naturally emerge as a consequence of the self-organizing nature of language.
- Functionalism in phonology is not teleological. It is the passive consequence of language use (Ohala…).
- The question pursued today could scarcely be formulated in a purely synchronic theory, let alone does synchronic theory provide the necessary tools to discover these patterns, these generalizations, and these answers.

Where to next?
- What did the root inventory of Middle Korean look like? (This is a very important issue!)
- What constitutes "significant homophony/ambiguity"?
- Look at other well known cases of neutralization/merger.
- Can we determine how much homophony/ambiguity a language can tolerate? Homophony? Realistically, only in terms of words, not tokens (and only at the lexical level, not at word boundaries).
- Ambiguity? No, we can't determine this.
- Can we undermine the notion that neutralization/merger is strongly counter-functional? Quite possibly…stay tuned…

Some references:
Albright, Adam. 2006. "Explaining universal tendencies and language particulars in analogical change," Manuscript, MIT.
Cho Seung-Bog. 1967. *A Phonological Study of Korean*. Uppsala: Almqvist and Wiksells.
Ito Chiyuki. 2006. "Coronal Coda Distribution in Korean Nouns," Manuscript, Tokyo University of Foreign Studies, ILCAA.

Jun Jongho. 2007. "Variation in Korean stem-final obstruents," Manuscript, SNU.

Kang Yoonjung. *To appear*. "Neutralizations and variations in Korean verbal paradigms," in Kuno, S. et al. eds., Harvard Studies in Korean Linguistics XI.

Kim Hyunsoon. 2001. "A phonetically based account of phonological stop assibilation," Phonology 18:81-108.

Kim-Renaud, Young-Key 1975. *Korean Consonantal Phonology*. Tower Press.

Lee Iksop, and Ramsey, S. Robert 2000. *The Korean Language*. Cambridge: Cambridge University Press.

Martin, Samuel E. 1992. *A Reference Grammar of Korean*. Rutland, VT: Charles E, Tuttle Company.

Sohn Ho-Minh. 1999. *The Korean Language*. Cambridge: Cambridge University Press.

Appendix

- There are other routes to neutralization and potential homophony. For example, here is an exhaustive list of noun homophones due to lateralization (n+ɾ, l+ɾ, l+n → [lː]; typos from the Sejong Project have been eliminated—Thanks to Jongho Jun for catching these).

Transliteration	Token Freq. (out of 1,353,602)	Hangul	Source	Gloss
{in-Ryu}	179	인류	SK	
{il-Ryu}	32	일류	SK	
{Keon-Rip}	55	건립	SK	
{Keol-Rip}	9	걸립	SK	
{Kweon-Ryeon}	2	권련	SK	
{Kweol-Ryeon}	10	궐련	SK	
{Pan-Ra}	2	반라	SK	
{Pal-Ra}	5	발라	SK	
{Pun-Ri}	124	분리	SK	
{Pul-Ri}	7	불리	SK	
{Pul-Ryang}	40	불량	SK	
{Pun-Ryang}	61	분량	SK	
{San-Rim}	35	산림	SK	
{Sal-Rim}	179	살림	SK	
{Sin-Rok}	3	신록	SK	
{Sil-Rok}	6	실록	SK	
{Sul-Nae}	4	술내	SK	
{Sul-Rae}	20	술래	SK	
{yeol-Rak}	2	열락	SK	
{yeon-Rak}	220	연락	SK	

- Ten (10) potential homophonous cases out of 1203 neutralizing forms (n+ɾ, l+ɾ, l+n → [lː]) in the list of 35,908 nouns.
- As for verbs, there are no potential homophonic sets along this parameter found in the database (304 neutralizing forms out of 3801 verbs, 644,403 tokens).
- And here is an exhaustive list of noun homophones due to nasalization (C_1+N → [N_1N] (C≠[l])).

Transliteration	Token Freq. (out of 1,353,602)	Hangul	Source	Gloss
{Cak-Mok}	7	작목	SK	
{Cang-Mok}	3	장목	SK	
{Cak-Mul}	35	작물	SK	
{Cang-Mul}	14	장물	SK	
{Cak-Mun}	24	작문	SK	
{Cang-Mun}	5	장문	SK	
{Cak-Nyeon}	267	작년	SK	
{Cang-Nyeon}	8	장년	SK	
{Chas-Mul}	2	찻물	SK	
{Chan-Mul}	26	찬물	SK	
{Hak-Mun}	234	학문	SK	
{Hang-Mun}	9	항문	SK	
{Kok-Mul}	20	곡물	SK	
{Kong-Mul}	3	공물	SK	
{Seok-Mul}	5	석물	SK	
{Seong-Mul}	17	성물	SK	
{yak-Mul}	42	약물	SK	
{yang-Mul}	7	양물	SK	

- Nine (9) potential homophonous cases out of 2109 neutralizing forms (C_1+N $\rightarrow$ [N_1 N] (C$\neq$[l]) in the list of 35,908 nouns (1,353,602 tokens).
- As for verbs, there is one (1) potentially homophonic set along this parameter found in the database (107 neutralizing forms out of 3801 verbs).

Transliteration	Token freq. (out of 644,403)	Hangul	Source	Gloss
{Man-Na}	1870	만나		
{Mas-Na}	7	맛나		