Analogical Changes in the Accent of Sino-Korean Words in Yanbian Korean

Keywords: analogical change, frequency, stratified lexicon, weighted constraints

This paper reports a statistical study of the Sino-Korean (SK) accent in Yanbian Korean (YK, spoken in north-eastern China and shows the factors working in its historical development. This is one of the results from the phonological study of YK which the author has been working on.

In Middle Korean (MK, 15th - 16th) where accent was distinctive, the accent of disyllabic SK words $\mu_1\mu_2$ was the combination of individual underlying accents of μ_1 and μ_2 as in (1).

(1)

μ_1	μ_2	$\mu_1\mu_2$	Examples
L	L	LL	kà 'house' + mùn 'gate' → kàmùn
L	H/R	LH	kà 'house' + sjók 'belonging' → kàsjók
H	ANY	HX	cjú 'main' + ìn 'person' → cjúin
R	ANY	RX	ɨɨ̆ 'discuss' + rŏn 'argue' → ɨɨ̆ron

L = low
H = high
R = rising
X = unspecified

The basic correspondence between MK and YK is straightforward: MK LL > YK LL, MK LH > YK LH, MK HX and RX > YK HL (merger). Our discussion focuses on patterned exceptions to these basic correspondences. Four major findings are reported.

First, a sharp distinction is drawn between native non-compound disyllabic nouns and SK disyllabic nouns. This is because native and SK words have different distributional patterns in the frequency of each accent class, as shown in Table 1 (next page). In SK words, HL is the biggest class (67%), followed by LH (29 %) and LL (4 %), whereas in native words LH is the biggest class (LH: 67 %, HL: 23 %, LL: 9 %). The correlation between analogical change and frequency has been discussed in many literatures (Hooper 1976, Bybee 1985, 2000, 2002, 2006, Phillips 1984, 2001, and others), where it has been pointed out that a higher type-frequency class tends to attract words from the lower type-frequency classes. In fact this is confirmed in our YK data, as shown in (2) and Table 2: in native words, where LH is the strongest class, LH tends to maintain the MK accent quite regularly (91 %) and rarely changed to the smaller class HL (only 5%), whereas smaller classes such as MK HX and RX often changed to LH irregularly (31 %) and the regular development rate was relatively lower (60%). In SK words, where HL is the strongest class, MK HX/RX quite regularly correspond with YK HL (79%) and the irregular development from these classes to LH is relatively rare (18%), whereas the smaller class LH does not show a high regular correspondence rate (52%) and it irregularly changed to the bigger class HL (= MK HX & RX) at a high rate (47%). Thus our first finding provides evidence for a **stratified lexicon**: speakers know to which word class (native or SK) the words belong and show different analogical changes depending on the type frequency within the word class.

(2)

	HX/RX → HL (regular)	HX/RX → LH (irregular)	LH → LH (regular)	LH → HL (irregular)
Native	60.19 %	31.07 %	91.32 %	4.96 %
SK	79.16 %	17.90 %	51.65 %	46.54 %

Second, although in SK HX/RX → HL is the most regular development, there is a striking exception where the accent of μ_2 determines a deviation to LH: when μ_2 corresponds with Middle Chinese (MC) Entering tone, which exclusively appears in the morphemes with codas -p/k/t (> SK -p/k/l), the accent of $\mu_1\mu_2$ tends to be LH except for the case which μ_1 also corresponds with Entering tone, as shown in Table 3 ("**Entering tone effect**," Island of Reliability for segmental phonology, Albright 2002). This is due to a high reliability in the correlation between Entering tone and H, which results from two factors: MC Entering tone regularly corresponded with MK H; Entering tone has a signature (coda -p/k/l) which identifies this class.

Given that the underlying accents of both μ_1 and μ_1 are playing roles in SK accent evolution, we propose a model with weighted constraints employing Jäger (to appear)'s Stochastic Gradient Ascent learning algorithm in which faithfulness constraints start at 0 and markedness at 10. The faithfulness constraints are F (E_1), F (H_1), F (R_1), F (L_1), F (E_2), F (H_2), F (R_2), F (L_2), where E means Entering tone class which appeared with H in MK, and the markedness constraints are *LL, *LH, *HL. The obtained weights are: *LL 11.26 > *LH 9. 84 > *HL 8.81 > F (E_1) 1.08 > F (L_1) 0.83 > F (H_2) 0.74 > F (E_2) 0.73 > F (R_2) 0.7 > F (H_1) 0.64 > F (L_2) 0.54 > F (R_1) 0.16. The fact that *HL is weakest among markedness constraints shows its default status as the largest group in SK. Given the default status of HL and that the higher weight for F (E_1) is due to the Island of Reliability effect for the syllable type, this weight ranking shows that the only information YK speakers have to memorize is underlying L which can result in LH or LL classes: in so far as speakers remember this information, the accent is almost automatically assigned to either HL or LH based on the markedness constraint ranking *LL > *LH > *HL. This is our third result.

Finally two factors play a role in deviations from the MK LH/LL to YK HL correspondence in SK words: when the onset of μ_1 is zero or sonorant, they correspond with YK LH/LL more regularly than when the onset is obstruent (Table 4, an asymmetry familiar from the tonogenesis literature); when μ_1 is a frequently used morpheme as the first element of disyllabic SK words, they correspond with YK LH/LL more regularly (Table 5). These correlations are statistically significant.

Table 1 Accent Distribution from a Yanbian native speaker (a female in her 30's)

	Native		SK	
HL	206	23.38 %	5368	67.10 %
LH	593	67.31 %	2301	28.76 %
LL	82	9.31 %	323	4.04 %
Totals	881		7992	

Table 2 Historical development (Regularity = regular development/total of the class)

Native

MK \ YK	HL	LH	LL	Totals	Regularity
HX	33	23	4	60	55.00
RX	29	9	5	43	67.44
LH	12	221	9	242	91.32
LL	10	60	33	103	32.04
Totals	84	313	51	448	70.54

SK

MK \ YK	HL	LH	LL	Totals	Regularity
HX	1603	243	40	1888	84.90
RX	1629	488	77	2195	74.21
LH	746	828	26	1603	51.65
LL	404	331	147	883	16.65
Totals	4382	1890	290	6569	64.04

Table 3 Detailed historical development in SK (E = Entering tone class)

MK \ YK	HL	LH	LL	Totals	Rate of LH
EE	298	43	8	349	12.32
EH	127	20		147	13.61
ER	461	66	8	535	12.34
EL	359	41	16	416	9.86
HE	93	28		121	23.14
HH	47	6		53	11.32
HR	110	20	1	131	15.27
HL	108	19	7	134	14.18

MK \ YK	HL	LH	LL	Totals	Rate of LH
RE	370	179	7	556	32.19
RH	183	18	1	202	8.91
RR	580	180	19	779	23.11
RL	496	111	50	657	16.89
LE	199	403	7	609	66.17
LH	149	111		260	42.69
LR	398	314	19	731	42.95
LL	404	331	147	882	37.53

Table 4 Correlation between accent change in MK LH/LL and onset of μ_1

ONSET \ YK	HL	LH/LL	Totals
Lax	522/ 486.56 (1.07)	526/ 561.44 (0.94)	1048
Aspirated	137/ 134.64 (1.02)	153/ 155.36 (0.98)	290
/h/	91/ 76.14 (1.20)	73/ 87.86 (0.83)	164
/s/	164/ 158.32 (1.04)	177/ 182.68 (0.97)	341
Ø	135 /173.64 (0.78)	239/ 200.36 (1.19)	374
Sonorant	101/ 120.71 (0.84)	159/ 139.29 (1.14)	260
Totals	1150	1327	2477

Left number = Observed, right number = Expected, ( ) = O/E values
$\chi^2 = 32.75$, $p = 4.23E-06$

Table 5 Correlation between accent change and type-frequency of μ_1 (p by Fisher's Exact Test)

Frequency	MK \ YK	HL	LH/LL	Totals	Regularity	p
high	E-	1022	166	1188	86.03	0.805
low	E-	365	56	421	86.70	
high	H-	300	64	364	82.42	0.192
low	H-	99	30	129	76.74	
high	L-	830	1071	1901	56.34	8.07E-09
low	L-	460	366	826	44.31	
high	R-	1290	413	1703	75.75	0.007
low	R-	513	215	728	70.47	

High when μ_1 occurs in more than 5 SK words.
'-' indicates any accent.

Selected References

Albright, Adam. (2002). Islands of reliability for regular morphology: evidence from Italian. *Language 78-4*, 684-709.

Bybee, Joan. (2006). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

Jäger, Gerhard. (to appear). Maximum Entropy Models and Stochastic Optimality Theory. In J. Grimshaw, J. Maling, C. Manning, J. Simpson, and A. Zaenen (Eds.), *Architectures, Rules, and Preferences: A Festschrift for Joan Bresnan*. Stanford: CSLI Publications.

Phillips, Betty. (2001). Lexical diffusion, lexical frequency, and lexical analysis. In Joan Bybee & Paul Hooper (Eds.), *Frequency and the Emergence of Linguistic Structure*: 123-136. Philadelphia: Benjamin.