Morphologically-based phonological variation

Jongho Jun (Seoul National University)

The present study explores the probabilistic application of a morpho-phonological process. A morpho-phonological process is typically conditioned by phonological and morphological factors. Such linguistic factors categorically distinguish words that can undergo the process from those that cannot. In contrast, non-linguistic factors such as word frequency typically have a gradient effect. They affect the process only probabilistically within the set of potential target words. However, some recent research (notably Zuraw 2010) shows that phonological and morphological factors may have a gradient effect, contributing to the overall probability of the occurrence of a morpho-phonological process. The present study explores the questions of what factors may have such a gradient effect in application of a morpho-phonological process, and how they interact, by investigating patterns of Korean n-insertion.

In Korean, /n/ is optionally inserted at the juncture of two morphemes, M_1 and M_2, under the condition that M_1 ends with a consonant and M_2 begins with a high front vocoid /i, j/:

(1) $\varnothing \rightarrow$ n / C_1]$_{M1}$ __ [$_{M2}$ i/j ($M_{1,2}$ = morpheme; C_1 = M_1-final consonant)
a. /com+jak/ [comnjak] ‘mothball’ b. /som+ipul/ [somnipul] ‘cotton sheet’

I conducted surveys on existing and novel Korean words, employing speakers of two dialects of Korean, Seoul and Kyungsang. From the results of the survey on 293 existing words with /j/-initial M_2, the following tendencies emerge:

(2) Tendencies in existing words: n-insertion is less likely ….

a. after a (bound) **root** than after a (free) stem.	(M_1 MORPHOLOGY EFFECT)
b. before a **root** than before a stem.	(M_2 MORPHOLOGY EFFECT)
c. after **Sino-Korean** morphemes than after native ones.	(M_1 ETYMOLOGY EFFECT)
d. before **Sino-Korean** morphemes than before native ones.	(M_2 ETYMOLOGY EFFECT)
e. after **monosyllabic** M_1 than after polysyllabic M_1.	(LENGTH EFFECT)
f. after an **obstruent** than a sonorant consonant.	(SONORANCY EFFECT)
g. after the **velar nasal** than other sonorant consonants.	(SONORANT PLACE EFFECT)
h. before **non-high** V_2 vowels than high V_2 vowels (V_2 = a vowel immediately following M_2-initial /j/)	(HEIGHT EFFECT)

What is more interesting is that some of the above tendencies vary depending on the morphological category of M_1 morphemes, as summarized below:

(3) Tendencies varying with morphology of M_1
a. SONORANT PLACE EFFECT is stronger for words with a free **stem** M_1 than a bound root M_1.
b. SONORANCY EFFECT is stronger for words with a **root** M_1 than a stem M_1.

I found a constraint weighting to explain the distribution of n-insertion in the existing word data, using a maxent (maximum entropy) learner implemented in the maxent grammar tool (Hayes 2009). The results of the existing word survey were employed as the input to the simulation, and the simulations were conducted separately for Seoul and Kyungsang data. Constraints adopted in the simulation and their learned weights are shown below:

(4) Constraint weights obtained using maxent grammar tool

	Seoul	Kyungsang	Constraint definition
INSERT-/n/	6.032	6.056	Insert /n/ between M_1 and M_2, if M_1 ends with a C and M_2 begins with /i j/.
DEP-n	3.968	3.944	Avoid insertion of /n/.
*INSERTION/ROOTM_1	0.689	0.478	If M_1 is a bound root, n-insertion is blocked.
*INSERTION/ROOTM_2	0.566	0.381	If M_2 is a bound root, n-insertion is blocked.
*INSERTION/SINOM_1	0.457	0.201	If M_1 is Sino-Korean, n-insertion is blocked.
*INSERTION/SINOM_2	0.521	0.253	If M_2 is Sino-Korean, n-insertion is blocked.
*INSERTION/MONOM_1	1.203	0.921	If M_1 is mono-syllabic, n-insertion is blocked.
IDENT(SON)/C_1	0.438	1.003	Input and output correspondents of C_1 have the same specification for [sonorant].
*INSERTION/ŋ_	0.389	0.478	If M_1 ends with /ŋ/, n-insertion is blocked.
*INSERTION/_j[-high]	0.469	0.461	If M_2 begins with /j[-high]/, n-insertion is blocked.

All the significant tendencies of existing words, summarized in (2, 3), are reflected in the learned weights of the corresponding constraints. The correlations between observed and predicted rates of n-insertion of the test words are relatively high and significant: Seoul ($r = 0.679$, $p < 0.001$, $n = 293$); Kyungsang ($r = 0.689$, $p < 0.001$, $n = 293$).

Finally, for the purpose of finding out whether Korean speakers are aware of all the tendencies observed in existing words, I explored n-insertion in novel Korean words by investigating the results of the surveys on native speakers of two dialects of Korean, Seoul and Kyungsang. Results show that SONORANT PLACE and HEIGHT EFFECTS are significant in novel words, being consistent with the results of an existing word survey. In contrast, SONORANCY and LENGTH EFFECTS are not significant in novel words. Given that M_1 morphemes used in the novel word survey were loan words, and should be interpreted as free stems, the insignificance of SONORANCY effect in the novel word survey can be interpreted to be consistent with the finding, reported above in (3b), that the SONORANCY effect was less prominent for existing words with a stem M_1. Consequently, Korean speakers are well aware of most gradient tendencies in the lexicon, in particular, those involving a morphology-phonology interaction.

References

Hayes, Bruce (2009). Maxent grammar tool, software, http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool.

Zuraw, Kie (2010) A model of lexical variation and the grammar with application to Tagalog nasal substitution. *NLLT* 28.2: 417-472.