How tight is the link between alternations and phonotactics?

Jongho Jun, Hanyoung Byun, Seon Park and Yoona Yee
Seoul National University

Supplementary materials

---

1. Details about the stimuli described in section 3.1.1 of the paper
2. Appendix: human and model ratings and relevant phenomena for test nonwords

REFERENCE
Jun, Jongho, Hanyoung Byun, Seon Park & Yoona Yee (2025). How tight is the link between alternations and phonotactics? *Phonology*.

1. Details about the stimuli described in section 3.1.1 of the paper

The test stimuli consisted of three sets, each of which was intended to investigate phonotactics related to one of the three effects (palatalisation, vowel harmony and laryngeal OCP). We did our best to ensure several conditions. First, the words do not sound like any particular real Korean words or phrases. They also do not violate any major phonotactic constraints of Korean other than those we were examining in this study. Further, all stimuli consist of two open syllables, of which the second may include a glide. Thus, the stimuli can be represented by $C_1V_1C_2(G)V_2$ where C = consonant, V = vowel and G = glide. However, three nonwords end with a consonant /n/. These exceptions were unavoidable because it was impossible for us to create a sufficient number of nonwords which obey all the relevant conditions required simultaneously. For V_1 and V_2 vowels, we use six simple vowels: /i e a o ʌ u/. We have excluded /ɛ/ which, as noted in section 2.2 of the paper, is not distinct from /e/ in younger speakers' speech, and /ɨ y ø/ which have limited distributions in Korean. C_1 and C_2 are never identical, controlling for segmental OCP effects. Finally, we counterbalanced the test stimuli as much as possible across consonant places and vowel sequences within each test set. Specific details about each set are provided below.

1.1 Palatalisation set

In this set, the lenis and aspirated coronal stops, /t t^h/, and their non-coronal counterparts, /p k p^h k^h/, occur in C_2 position. Tense stops were excluded from C_2 position since no Korean morphemes end in a tense coronal stop /t'/, and thus there are no attested cases in which tense stops undergo palatalisation in Korean. The high front vocoids, /i j/, and the other glide, /w/, occur in V_2 and G positions. As shown in (20) of the paper, 20 target nonwords with TI sequences and 20 control nonwords with non-coronal stops followed by high front vocoids such as /pi/ (henceforth PI) were created. In addition, control nonwords with coronal stops followed by vocoids other than high front vocoids such as /ta/ (henceforth TA) were created. Specifically, 10 control nonwords with vowels other than /i/ and 4 control nonwords with /w/ were added to the list. For these items, five vowel pairs (a-o, o-a, a-u, i-u, ʌ-u), which are frequent in our Seoul Korean corpus (see section 4 of the paper), were used for V_1-V_2. The reason why 4, not 10, control nonwords with /tw t^hw/ were created in the experiment is that /w/ in Korean is subject to some distributional restrictions, especially on its combination with non-low back vowels, (Kwon 2018) and thus it was impossible to create a sufficient number of novel words involving /tw t^hw/. For the positions which were not relevant for the current purpose, we used /p c k m n/ in C_1 and /a o ʌ u/ in V_1. In addition, for the items that have G, we used /a ʌ/ in V_2 position.

Six nonwords with non-coronal stops followed by a labial-velar glide such as /pw/ were adopted to reduce the imbalance of the stimuli set in which a palatal glide /j/ was used predominantly over a labial-velar glide /w/. In addition, six nonwords, which were originally created to investigate laryngeal OCP and/or vowel harmony, turned out to be useful to investigate palatalisation, meeting all the relevant conditions (such as having no violations of known phonotactic constraints other than palatalisation, having non-identical consonants for C_1-C_2 sequences, and having non-identical vowels in V_1-V_2 sequences), and thus were also considered in the analysis as part of the current set of nonwords. The six additional items had the following C_2-V_2 vowel sequences (the number of nonwords of corresponding type is shown in parentheses): tu (1), t^hu

(2), pi (2) and pʰi (1).

1.2 Vowel harmony set

In this set, the V_1-V_2 sequences utilised exhausted all possible sequence permutations of the six vowels, /a e i o u ʌ/, and for each V_1-V_2 sequence, two nonwords were created, thus yielding 72 nonwords in total (6 vowels x 6 vowels x 2 items), as shown in (21) of the paper. Eight nonwords, which were originally created to investigate palatalisation, turned out to be useful to investigate vowel harmony, meeting the relevant conditions (such as having no violations of known phonotactic constraints other than vowel harmony), and thus were also considered in the analysis as part of the relevant set of nonwords. The eight additional items had the following V_1-V_2 sequences: a-o, a-u, i-u, o-a, u-i, ʌ-u and ʌ-i. Two items with ʌ-i and one item with each of the rest sequences were included.

1.3 Laryngeal OCP set

In this set, the C_1-C_2 onset sequences utilised exhausted all possible sequence permutations of three consonant phonation types, i.e., lenis, aspirated and tense. For each C_1-C_2 phonation type sequence except for lenis-lenis, nonwords with eight different V_1-V_2 sequences were created. The following V_1-V_2 sequences, which are frequent in our Seoul Korean corpus, were used: /a-i/, /u-i/, /a-o/ /u-o/, /ʌ-u/, /a-u/, /i-u/ and /o-a/. Eight lenis-lenis nonwords had already been created for the other two conditions, thus making up 72 nonwords (3 phonations x 3 phonations x 8 vowel sequences), as shown in (22) of the paper. For C_1 and C_2 consonants, we used four lenis obstruents /p t c k/ as well as their aspirated and tense counterparts. /s s'/ were excluded because they lack an aspirated counterpart in Korean, as can be seen in (14) of the paper. Among the nonwords created to investigate vowel harmony and/or palatalisation, eleven of them, including eight items with lenis-lenis sequences, met the relevant conditions required for this set of nonwords, such as having no violations of known phonotactic restrictions other than the laryngeal OCP and having vowel sequences consistent with those adopted for this set mentioned above. The remaining three items have lenis-aspirated sequences. These items were thus also considered in the analysis as part of this set of nonwords.

2. Appendix: human and model ratings and relevant phenomena for test nonwords

stimulus	human rating	model rating	palatalisation	vowel harmony	laryngeal OCP
pʰ u t o	2.45	0			o
tʰ u p i	2.69	0			o
cʰ ʌ k u	4.35	0			o
kʰ a c i	2.33	-2.757			o
pʰ a k u	4.2	0			o
tʰ o c a	3.51	0			o
cʰ i t u	3.96	-1.621			o
kʰ a p o	2.35	-2.757			o
p o cʰ a	3.59	0			o
t a pʰ i	2.9	0	o		o
c u kʰ o	2.96	-2.757			o
k i tʰ u	3.65	-1.918	o		o
p ʌ tʰ u	2.86	-3.29	o		o
t a kʰ o	2.76	-2.757			o
c u pʰ i	3.14	-1.631	o		o
k a cʰ u	3.84	0			o

p’ a t u	1.71	-3.962			o
t’ a p o	1.84	-1.993			o
c’ o k a	2.39	-3.872			o
k’ ʌ c u	2.45	-1.993			o
p’ u c i	1.96	-3.962			o
t’ u k o	2.06	-1.993			o
c’ a p i	2.27	-3.872			o
k’ i t u	2.04	-3.615			o
p o c’ a	2.18	-3.249			o
t a p’ o	3.47	-3.249			o
c a k’ i	3.2	-3.249			o
k a t’ u	2.04	-3.249			o
p i t’ u	2.08	-6.789			o
t u k’ o	2.53	-3.249			o
c u p’ i	2.22	-3.249			o
k ʌ c’ u	2.27	-4.831			o
p^h u c^h i	1.73	-1.256			o
t^h a p^h u	1.92	-1.256			o
c^h a k^h o	3.14	-4.013			o
k^h u t^h o	1.84	-5.702			o
p^h o t^h a	2.86	-1.256			o
t^h a k^h i	1.78	-4.013			o
c^h ʌ p^h u	2.35	-2.937			o
k^h i c^h u	2.18	-4.013			o
p^h i t’ u	1.49	-7.353			o
t^h u k’ o	2.39	-3.813			o
c^h ʌ p’ u	3.61	-5.494			o
k^h a c’ i	1.67	-6.57			o
p^h u c’ i	1.71	-3.813			o
t^h a p’ o	2.08	-3.813			o
c^h o k’ a	3.39	-3.813			o
k^h a t’ u	1.53	-6.57			o
p’ o t^h a	1.67	-5.218			o
t’ a k^h u	2.18	-6.006			o
c’ u p^h i	1.94	-6.759			o
k’ i c^h u	1.96	-3.25			o
p’ u c^h o	1.49	-5.218			o
t’ a p^h i	1.92	-3.25			o
c’ ʌ k^h u	1.78	-7.885			o
k’ a t^h o	1.98	-4.874			o
p’ o c’ a	1.67	-7.211			o
t’ a p’ i	1.55	-5.243			o

c’ u k’ o	1.94	-7.121			o
k’ ʌ t’ u	1.67	-5.243			o
p’ a t’ o	1.35	-7.211			o
t’ a k’ u	2.41	-5.243			o
c’ i p’ u	1.86	-7.121			o
k’ u c’ i	1.65	-5.243			o
p a n a	4.12	0		o	
t ʌ m a	3.41	0		o	
k o n a	4.1	0		o	
p u t a	4.1	0		o	
n e p a	3.12	0		o	
m i t a	4.18	-1.621		o	
t a p a	4.04	0		o	
p ʌ k a	3.35	0		o	
t o p a	3.96	0		o	o
k u p a	3.37	0		o	
c e m a	4.73	0		o	
m i k a	4.8	0		o	
n a p ʌ	3.12	0		o	
k ʌ p ʌ	3.29	0		o	
c o k ʌ	3.65	0		o	
p u n ʌ	3.73	-1.77		o	
t e p ʌ	3.94	0		o	
k i m ʌ	2.86	-2.675		o	
c a m ʌ	3.47	-2.675		o	
m ʌ t ʌ	2.78	-2.675		o	
n o k ʌ	4.24	0		o	
m u c ʌ	4.18	0		o	
n e p ʌ	2.35	0		o	
p i t ʌ	3.22	-1.621		o	
t a m o	4.67	0		o	
t ʌ p o	3.84	0		o	
k o n o	3.35	0		o	
k u n o	4.02	0		o	
t e n o	5.2	-2.419		o	
k i n o	2.98	0		o	
k a m o	4.53	0		o	
c ʌ k o	4.63	0		o	
m o t o	3.96	0		o	
t u n o	3.53	0		o	
p e c o	4.49	0		o	
p i n o	3.73	0		o	

t a m u	4.45	0		o	
p ʌ n u	3.65	-1.692		o	
t o m u	4.8	0		o	
n u p u	4.2	-1.692		o	
k e m u	4.78	0		o	
c i m u	4.08	0		o	
p a c u	4.14	0		o	o
n ʌ k u	4.08	-1.77		o	
p o n u	3.71	-1.692		o	
p u k u	4.35	0		o	
m e t u	4.41	0		o	
p i t u	4.2	-1.621	o	o	o
c a p e	3.94	0		o	
m ʌ t e	3.63	-2.675		o	
n o p e	3.82	0		o	
k u t e	3.57	0		o	
k e m e	3.29	0		o	
p i m e	3.98	0		o	
m a p e	3.92	0		o	
t ʌ p e	3.29	0		o	
m o c e	4.49	0		o	
n u p e	3.98	-1.692		o	
t e m e	3.63	0		o	
k i p e	4.27	0		o	
t a k i n	4.02	0		o	
m ʌ p i	3.29	-2.675	o	o	
m o p i	4.12	0	o	o	
c u p i	3.63	0	o	o	o
p e n i	3.02	-2.419		o	
p i m i	3.82	0		o	
k a m i n	4.9	0		o	
t ʌ p i	3.39	0	o	o	
c o n i n	3.06	0		o	
t u m i	4.49	0		o	
k e n i	2.96	-2.419		o	
k i p i	3.55	0		o	
m ʌ t i	2.9	-5.96	o		
k o t i	4.04	-3.286	o		
n a t i	3.49	-3.286	o		
k ʌ t i	3.12	-3.286	o		
c a t i	2.96	-3.286	o		
p a t o	4.35	0	o	o	o

n o t a	4.18	0	o	o	
p a t u	4.24	0	o	o	o
k i t u	4.71	-1.621	o	o	o
m ʌ t u	4.06	-2.675	o	o	
p o tʰ i	2.31	-3.286	o		
n a tʰ i	2.1	-4.91	o		
m ʌ tʰ i	2.49	-9.251	o		
c u tʰ i	2.18	-4.975	o		
k a tʰ i	3.12	-4.91	o		
p a tʰ o	3	-1.624	o		o
k o tʰ a	3.27	0	o		o
n a tʰ u	2.61	-1.624	o		
c i tʰ u	3.1	-1.918	o		o
m ʌ tʰ u	2.92	-5.965	o		
t u p i	3.9	0	o	o	o
n ʌ k i	3.67	-1.77	o	o	
p ʌ k i	3.24	0	o	o	
n a pʰ i	3.2	0	o		
t o kʰ i	2.8	-2.757	o		
m o kʰ i	2.94	-2.757	o		
p ʌ kʰ i	1.86	-2.757	o		
k ʌ t j a	1.96	-12.64	o		
n o t j a	1.76	-12.64	o		
p a t j ʌ	2.12	-12.64	o		
m u t j a	2.14	-12.64	o		
c a t j ʌ	1.92	-12.64	o		
m ʌ tʰ j a	2.04	-14.889	o		
p o tʰ j ʌ	2.18	-8.924	o		
k a tʰ j ʌ	2.49	-10.548	o		
c u tʰ j a	1.69	-10.613	o		
n a tʰ j ʌ	1.94	-10.548	o		
m u t w a	2.08	-2.891	o		
c u t w ʌ	2.59	-2.891	o		
n ʌ k w a	3.59	-1.77	o		
p a k w ʌ	2.69	0	o		
k o p w ʌ	2.39	-2.98	o		
p o tʰ w a	2	0	o		
m a tʰ w ʌ	2.33	-1.624	o		
n ʌ kʰ w a	3.47	-4.526	o		
c u pʰ w a	2.18	-3.954	o		
k a pʰ w ʌ	2.65	-2.323	o		
n ʌ p j a	2.29	-1.77	o		

k o p j ʌ	3	0	o		
c a k j ʌ	4.57	0	o		
m o k j a	2.78	0	o		
p u k j ʌ	3.53	0	o		
c ʌ p^h j a	2.14	-1.68	o		
n a p^h j ʌ	2.45	0	o		
k o p^h j a	1.94	-1.631	o		
m ʌ k^h j a	1.94	-5.431	o		
p u k^h j ʌ	2.45	-2.757	o		

References

Kwon, Soohyun (2018) *The Development of Glide Deletion in Seoul Korean: A Corpus and Articulatory study.* Ph.D. dissertation, University of Pennsylvania.