

## Linguistic entrenchment and the effect of subjective lexical familiarity in Korean /n/-insertion

Jiyeon Song & Amanda Dalola\*

**Abstract.** Korean shows variable /n/-insertion between a morpheme-final consonant and the initial /i/ or /j/ of a following morpheme. Literature has shown that the appearance of the phenomenon can be affected by various parameters, including social and phonological factors. Exemplar theory contends that a word’s susceptibility to language variation correlates directly with its word frequency, a unitary frequency measure based on a corpus (Pierrehumbert 2001; Bybee 2002). However, given that individuals have different language experience, word frequency rarely addresses individual differences in the same way that self-rated measures of word frequency, known as subjective lexical familiarity, do. This research investigates whether and how the metric of self-rated lexical familiarity affects Korean /n/-insertion. Results indicate that subjective lexical familiarity significantly predicts the appearance of /n/-insertion, such that words more familiar to the speaker show /n/-insertion more often than those that are less familiar.

**Keywords.** Korean /n/-insertion; subjective lexical familiarity; variation; sociophonetic factors

**1. Introduction.** In Korean, /n/-insertion appears variably when a first morpheme is consonant-final and a second morpheme begins with a high front vocoid, /i/ or /j/. Literature on the phenomenon suggests that it is affected by various sociolinguistic, phonological, and morpho-phonological factors, such as speaker age, dialect, historical source of the word (native vs. Sino-Korean), historical source of the first morpheme-final consonant, manner of the first morpheme-final consonant, vocoid type and speaker age-specific lexical familiarity (or lexical familiarity regarding generation) (Kim 2013, Kim 2003, Kook et al. 2005, Song & Dalola 2018). This research aims to investigate the conditioning effect of seven social, phonological and historical factors on /n/-insertion, including the novel variable of subjective lexical familiarity (SLF).

### 2. Review of literature.

2.1. KOREAN /n/-INSERTION. Korean shows /n/-insertion at the juncture of two morphemes comprising a compound (1a), a derivational word (1b), or a syntactic phrase (1c), especially when the first morpheme ends in a consonant (C<sub>1</sub>) and the second morpheme begins with a vocoid, the high front vowel /i/ or the glide /j/, as in (1).

(1)	underlying form	no /n/-insertion (option Y)	/n/-insertion (option Z)	morphological category	gloss
a.	/som-i.pul/	[so.mi.pul]	[som.ni.pul]	compound	‘cotton comforter’
b.	/sɛk-jʌn.p <sup>h</sup> il/	[sɛ.kjʌn.p <sup>h</sup> il]	[sɛŋ.njʌn.p <sup>h</sup> il]	prefix-stem	‘colored pencil’
c.	/os-ib.ta/	[o.tib.ta]	[on.nip.ta]	noun-verb	‘put on clothes’

\* Special thanks to Taebeom Lee, Kyungsook Kim, and Kyongsok Lee for data collection.

Authors: Jiyeon Song, University of South Carolina ([jiyeons@email.sc.edu](mailto:jiyeons@email.sc.edu)) & Amanda Dalola, University of South Carolina ([dalola@mailbox.sc.edu](mailto:dalola@mailbox.sc.edu)).

There is a variant of /n/-insertion in which the nasal /n/ does not appear and instead the resyllabification of a morpheme-final consonant takes place, as in option Y in (1). Both /n/-insertion (option Z) and the variant without /n/-insertion (option Y) produce acceptable surface variants, as depicted in (1).

2.2. MOTIVATION OF /n/-INSERTION. Table 1 shows the Korean consonant chart. Korean has several consonants that are less marked than the nasal /n/ which it might use for epenthesis (e.g., /h/ and /t/) (Vaux 2002).

		Bilabial	Alveolar	Palatal	Velar	Glottals
Stops (plosives)	Lenis	p	t		k	
	Aspirate	p <sup>h</sup>	t <sup>h</sup>		k <sup>h</sup>	
	Fortis	P*	tt		k*	
Fricatives	Lenis		s			h
	Fortis		s*			
Affricates	Lenis			tɕ		
	Aspirate			tɕ <sup>h</sup>		
	Fortis			tɕ*		
Nanals		m	n		ŋ (syllable-final)	
Laterals			l			

Table 1. Korean consonants (Shin & Cha 2003)

The insertion of a relatively marked nasal by the process of Korean /n/-insertion begs the question of why the coronal nasal /n/ is inserted but not other coronals. Kim et al. (2007) found evidence for a phonetic motivation: they compared the first formant (F1) and the second formant (F2) transition of /i/ and after /n/ and /l/. In their estimation, /n/ and /l/ are “legitimate candidates” because they produce continuing sonority across a syllable boundary (also known as Syllable Contact Law) and have “phonetic similarity to vowels” (Kim et al. 2007: 194). Their results show that /n/ is the less salient consonant before /i/ than /l/. That is, the F1 and F2 transition of /l/ to the following vowel /i/ shows more radical change than the transition between /n/ and /i/. The nasal /n/ can, thus, be viewed as epenthetically ideal for the smooth transition in sonority it offers between the morpheme-final consonant to the morpheme-initial vowel.

According to Kim et al. (2007), if there is no /n/-insertion, radical transition from C<sub>1</sub> to the following vocoids will occur in the formants and the sonority slope will also starkly rise. Thus, resyllabification without /n/-insertion may be viewed a less preferred strategy than /n/-insertion. However, since free variation between /n/-insertion and resyllabification is readily attested, it is apparent that salience is not sufficient to explain the phenomenon. Thus, this study will probe the phonological and social predictors conditioning this variation.

2.3. SOCIAL FACTORS: AGE AND DIALECTS. Previous work has argued that age and dialect type play an important role in the variation of Korean /n/-insertion. Kook et al. (2005) tested whether /n/-insertion preferences have changed over time, to be indicative of a generation gap. They conducted a survey with 183 words in the top 1500 most frequent words of the language. The words included 57 native-Korean words and 126 Sino-Korean words. 167 natives of Seoul Korean (SK) were instructed to choose their preferred pronunciation from two to four options. The results show that /n/-insertion preferences are affected by the speakers’ age such that participants be-

tween their 50s and 70s were more likely to choose pronunciations including /n/-insertion than those in their 20s to 40s.

/n/-insertion also shows variation depending on dialect. Kim (2013) investigated the presence of /n/-insertion in Sino-Korean. She conducted a production task with 17 NKK natives and 18 SK natives and found that both dialects showed different tendencies depending on the morphological categories of words involved. NKK speakers showed overwhelmingly more /n/-insertion than SK speakers in Sino-Korean root compound words (combination of two roots) (66.7% vs. 0.2%). However, there was no significant difference across dialects for Sino-Korean compounds (combination of two words) (NKK: 87.1% vs. SK: 76.7%). It is, therefore, imperative to further the investigation of /n/-insertion in the two dialects by applying inferential statistics.

2.4. PHONOLOGICAL AND HISTORICAL FACTORS. Several phonological factors have been examined with respect to Korean /n/-insertion, including the phonological features of  $C_1$ , the nature of morpheme-initial vocoids, and the historical source of words. Kook et al.'s (2005) work reveals that manner of  $C_1$  and historical sources of words do significantly affect the occurrence of /n/-insertion. Sino-Korean words show /n/-insertion more than native-Korean words, and both Sino- and native-Korean words show different preferences for /n/-insertion depending on the identity of  $C_1$ , in the following order: /m/, /n/, /l/, /ŋ/. As for vocoids, /n/-insertion is found to be more likely to appear before the glide /j/ than the vowel /i/ in both Sino-Korean and native Korean words. Similarly, Um (2010) has suggested that /n/-insertion may be disappearing gradually, especially in words containing the vowel /i/ as the morpheme-initial vocoid.

Hwang (2008) conducted a wug test on 20 SK natives (11 females, 9 males) to investigate which factors influence /n/-insertion. Participants were required to read aloud the stimuli in a sentence. Presence/absence of /n/-insertion was evaluated with acoustic analysis through spectrogram inspection in Praat. The results reported that /n/-insertion is more likely before /j/ than before /i/, and that it was less likely after obstruents than after sonorants other than /ŋ/.

Jun (2015) explores patterns of /n/-insertion with current SK words from three studies: Kang & Kim (2004), Choi (2002), and Kim (2003). He carried out a survey with 22 SK speakers using 304 words including the glide /j/<sup>1</sup>, a vocoid in the second morpheme. The goal was to investigate patterns and distribution of /n/-insertion in current Seoul Korean words. For this, he constructed a dictionary database including words which were attested in the Sejong text corpus of 5.5 million words (National Institute of the Korean Language 2014). He excluded words that are very rarely used in “real speech situations” to focus on SK speakers’ production and intuition about /n/-insertion (Jun 2015: 423). The results of his survey indicate that /n/-insertion is significantly dispreferred when  $C_1$  is an obstruent or the velar nasal /ŋ/, as compared to other sonorants, however, it is significantly more frequent before /ju/ than before other vocoid and vowel sequences. /n/-insertion is also less likely to happen with a longer first morpheme than with a shorter first morpheme in terms of number of syllables. More interestingly, when the first morpheme (M1) is a Sino-Korean word, /n/-insertion is less likely to occur than when M1 is a native Korean word. Although the percentages between the two were 45.96% (native Korean) and 43.13% (Sino-Korean), this is the opposite tendency reported in previous research (Kook et al. 2005). The Standard Korean Dictionary (Kwuklip kwuke yenkwuwen 1999) includes 25.28%

---

<sup>1</sup> Jun (2015) argues that multimorphemic words including the glide /j/ as a vocoid of the second morpheme (M2) frequently show /n/-insertion, and thus focused specifically on these words, not on those including vowel /i/.

native-Korean words, 57.12% Sino-Korean words, 5.26% loanwords, and 12.28% words of mixed historical sources (Chung 2000:63). Considering the different proportion of Sino- and native-Korean words, and the conflicting results from previous research (Kook et al. 2005 and Jun 2015), it is imperative to verify the effect of historical source using inferential statistics.

2.5. SUBJECTIVE LEXICAL FAMILIARITY. Traditional generative models of phonology and phonetics cannot explain why phonetic variation “appl[ies] differently to some specific words than to others” (Pierrehumbert 2000:2). The usage-based framework has emerged to resolve the issue with word frequency. Exemplar theory posits that phonological patterns and variations can be learned through “experience with speech” (Pierrehumbert 2000: 1). However, standard “objective” language-level word frequency does not reflect individuals’ actual exposure to words and their rates of use in varying social and pragmatic situations. Rather, it shows a possibility of encounter with words based on a specific arbitrarily assembled corpus. Kook, Kim & Lee (2005) report that agricultural terms that are unfamiliar to a younger generation exhibit much less /n/-insertion among younger speakers than in an older generation. Although they use words in the top 1500 most frequent words, each individual’s subjective lexical familiarity (SLF) may uniquely condition the presence of /n/-insertion. For this reason, the present study will make use of self-rated measures of individual word familiarity in order to examine whether subjects’ personal experience with words might affect their n/-insertion behaviors.

### 3. Research questions and variables

3.1. RESEARCH QUESTIONS. The focus of our inquiry is to investigate the presence of inserted /n/ between two morphemes, as conditioned by SLF, social factors (dialect, gender, and age), phonological factors (manner of C<sub>1</sub> and nature of vocoid) and historical origin. This research proposes four main research questions: i) Does SLF affect /n/-insertion? ii) What social factors (dialect, gender, and age) affect Korean /n/-insertion? iii) What phonological factors (manner of C<sub>1</sub> and nature of vocoid) affect /n/-insertion in Korean? iv) Does historical origin (native vs. Sino-Korean) affect /n/-insertion?

3.2. VARIABLES. In this study, the dependent variable is the binary presence/absence of /n/-insertion. There are seven categorical independent variables, operationalized with the following levels: dialect (Seoul, Kyungsang), gender (male, female), age (20s, 60s), historical source (native, Sino-Korean), manner of C<sub>1</sub> (stop, nasal, approximant, fricative, affricate), nature of vocoid (/i/, /j/), and SLF (high, intermediate, low).

### 4. Methodology.

4.1. PARTICIPANTS. 20 native NKK speakers and 20 native SK speakers balanced for gender from two different generation groups (young=20s, older=60s) were recruited via their indication of dialects/parent dialects on a self-reporting survey. All participants and their parents were born and grew up in their reported speech communities.

4.2. STIMULI. The wordlist consisted of 53 words ranging from two to four syllables in length. 48 nouns and five verbs were included. As the focus of this study was on phonological variation, words with obligatory /n/-insertion were excluded<sup>2</sup>. The words were presented in Korean Hangul

---

<sup>2</sup> According to Song & Dalola (2018), when C<sub>1</sub> is a trace of a genitive case marker from Middle Korean, it is highly possible that resyllabification is not allowed (e.g. /namunnip/ ‘leaf’ and /pekennit/ ‘pillowcase’). Additionally, some verb phrases including sentence ender ‘ta’ do not allow /n/-insertion (e.g. /masitt’a/ ‘delicious’ and /mæsitt’a/) (Ko 1992:32).

script. All words in the list do not represent their range of surface pronunciations in the writing system. That is, they do not visually show the two possible pronunciations for /n/-insertion and resyllabification; they only show orthographic representation, which does not overtly express /n/.

4.3. PROCEDURE. Participants completed a wordlist reading task. They were instructed to read the script naturally in a way deemed acceptable by their speech community. Recordings were made on a Marantz Professional PMD661 MKIII, digitized at a 44100 Hz with 16-bit accuracy.

4.4. MEASUREMENT. Presence/absence of /n/-insertion was determined via auditory observation and the identification of nasal formants arising from nasal resonance (Boersma, 2001). Each of the words was examined for the presence of the nasal /n/ using the waveform and the spectrogram, as in Figure 1.

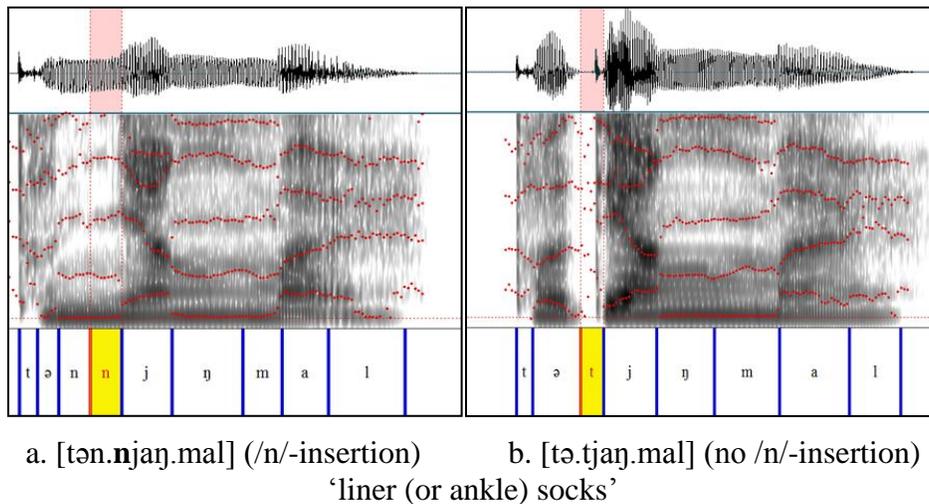


Figure 1. Absence and presence of /n/-insertion

**5. Results.** Table 2 illustrates the raw counts and percentages of /n/-insertion in seven variables: dialect, gender, generation (age), historical source, manner of C<sub>1</sub>, nature of vocoid and SLF.

n=2100	Yes /n/-insertion	total
Seoul	727 (68.58%)	1060
Kyungsang	746 (70.38%)	1060
male	772 (72.83%)	1060
female	701 (66.13%)	1060
younger generation (20s)	705 (66.51%)	1060
older generation (60s)	768 (72.45%)	1060
native Korean words	978 (62.69%)	1560
Sino-Korean words	495 (88.39%)	560
stop	524 (65.50%)	800
nasal	443 (92.29%)	480
approximant	157 (98.13%)	160
fricative	259 (49.81%)	520
affricate	90 (56.25%)	160
nature of vocoid /i/	584 (60.83%)	960
nature of vocoid /j/	889 (76.64%)	1160

SLF high	1121 (79.33%)	1413
SLF intermediate	217 (52.29%)	415
SLF low	135 (46.23%)	292

Table 2. Occurrences of /n/-insertion

A logistic regression was used to examine the role of the independent variables to condition the appearance of /n/-insertion. Statistical analyses were computed using the parameter *family=binomial* in the *glm()* function in the statistical tool R (version 3.4.1: R Core Team, 2007). The results are given in Table 3.

	Odds ratio	Std. Error	p-values
intercept	2.21	0.1351	4.36e-09 ***
mannerof C <sub>1</sub> nasal	4.44	0.1947	1.98e-14 ***
mannerof C <sub>1</sub> approximant	17.22	0.5914	1.49e-06 ***
mannerof C <sub>1</sub> fricative	0.56	0.1216	1.50e-06 ***
mannerof C <sub>1</sub> affricate	0.83	0.1831	0.3074
SLF intermediate	0.50	0.1290	3.55e-08 ***
SLF low	0.30	0.1502	1.10e-15 ***
nature of vocoid /j/	1.94	0.1121	3.65e-09 ***
gender female	0.66	0.1060	7.06e-05 ***
generation older (60s)	1.31	0.1062	0.0114 *

Note. . = .05 < p < .1, \* = p < .05, \*\* = p < .01, \*\*\* = p < .001

Table 3. Logistic regression for occurrence of /n/-insertion

Results found a main effect for C<sub>1</sub> (p < .001) such that morpheme-final nasals were 4.44 times more likely to show /n/-insertion than the reference group of stops, and morpheme-final approximants were 17.22 times more likely to show /n/-insertion than the reference group of stops. There was also a main effect for C<sub>1</sub> fricatives (p < .001) such that morpheme-final fricatives were 44% less likely to show /n/-insertion than the reference group of stops. A main effect for SLF emerged (p < .001), such that words with intermediate SLF were 50% less likely to show /n/-insertion than the reference group of words with high SLF, and words with low SLF were 70% less likely to show /n/-insertion than the reference group of words with high SLF. Nature of vocoid shows a main effect (p < .001) such that words including the glide /j/ were 1.94 times more likely to show /n/-insertion than those including the high front vowel /i/. Gender also shows a main effect (p < .001), such that females were 34% less likely to show /n/-insertion than males. Additionally, the older generation in their 60s was 1.31 times more likely to show /n/-insertion than younger generation in their 20s (p < .05). There was no main effect for dialect or historical source. Significant interaction among independent variables was not observed.

**6. Discussion.** The results show five main effects on /n/-insertion variation. First, females are less likely to show /n/-insertion than males. Females have often been credited with leading sound changes in progress (Labov 1966; 1984). Supposing /n/-insertion is a sound change in progress, we may postulate that females are leading the way with respect to /n/-insertion. Secondly, the older generation in their 60s is more likely to show /n/-insertion than the younger generation in their 20s, a finding that is consistent with Kook et al. (2005). Thirdly, /n/-insertion is more likely to appear after nasals and approximants than stops, fricatives and affricates, corroborating Jun's

(2015) finding that sonorants condition /n/-insertion at higher rates than obstruents. Fourth, /n/-insertion is more likely to appear before the glide /j/ than before the vowel /i/, which is consistent with Kook et al. (2005). Fifth, words unfamiliar to the speaker show /n/-insertion less often than familiar words. This suggests that linguistic entrenchment conditioned by frequent exposure to certain forms may, in fact, cause /n/-insertion “to be processed as a holistic unit” and to resist the variation common in language change (Pagel et. al 2013; Schmid 2014). In this way, we find evidence to suggest that SLF takes into account morpho-phonological strategies that are reflective of both language-level frequency (i.e. word frequency) and individual-level frequency (i.e. subjective lexical familiarity). The results did not show a main effect for dialect or historical source. This partially supports Kim’s (2013) finding that dialect effects can be different depending on the morphological categories of words. Considering the inconsistent and descriptively similar results reported for the role of historical source on /n/-insertion (Kook et al. 2005, Jun 2015), the lack of a main effect in our model after the application of inferential statistics is unsurprising.

**7. Conclusion.** This research has found a main effect of SLF on /n/-insertion and provided evidence to suggest that SLF may be a more comprehensive metric of frequency in studies of variation and language change. According to the “non-stationary effect” identified by Hay et al. (2015), the effect that word frequency can have on language change is different according to the momentum exhibited by the sound change. It is, therefore, possible that the variability in Korean /n/-insertion may be indicative of long-term language change, whereby less frequent forms show a preference for resyllabification and more frequent forms show a preference for /n/-insertion. Thus, the entrenchment of high-frequency words resists replacement by resyllabification, while low-frequency words lead linguistic change via analogy (Hay et. al 2015). In this respect, Korean /n/-insertion illustrates the nuanced role fulfilled by both individual experiences and analogy in long-term language change.

## References

- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10). 341-345.
- Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14(3). 261-290. <https://doi.org/10.1017/S0954394502143018>.
- Choi, Hye-Won. 2002. *A survey of standard pronunciation*. [Phyocunpalum silthaycosa.] Seoul: National Institute of the Korean Language.
- Hay, Jennifer, Janet B. Pierrehumbert, Abby J. Walker & Patrick LaShell. 2015. Tracking word frequency effects through 130 years of sound change. *Cognition*, 139, 83-91. <https://doi.org/10.1016/j.cognition.2015.02.012>.
- Jeong, Hoseong. 2000. Statistical analysis of the Standard Korean Dictionary [Phyocwunkwuketaysacen swulokcengpouy thongkyeysek pwunsek]. *New Korean Language Life* 10(1). 55-72.
- Jun, Jongho. 2015. Korean n-insertion: a mismatch between data and learning. *Phonology* 32(3). 417-458. <https://doi.org/10.1017/S0952675715000275>
- Kang, Beom-mo & Kim Hung-gyu. 2004. *Frequency analysis of Korean morpheme and word usage*. [Hankwuke hyengtayso mich ehwi sayong pintouy pwunsek.] Vol. 2. Seoul: Institute of Korean Culture, Korea University.

- Kim, Hyung-Sun, Baegseung Kim & Mira Oh. 2007. An optimality theoretic analysis of phonetically motivated/n/-insertion. *The Linguistic Association of Korea Journal* 15. 187-205. <http://cm620.cafe24.com/data/thesis/1167131047511.pdf>
- Kim, J. 2013. *Differences between two dialects in /n/-insertion of Sino-Korean: Speakers in their 20s in Northern Kyungsang Korean and Seoul Korean*. Korea University. Master Dissertation.
- Kim, Seoncheol. 2003. *A survey of standard pronunciation*. [Phyocwunpalum silthaycosa.] Vol. 2. Seoul: National Institute of the Korean Language.
- Ko, Kwang-mo (1992). n-epenthesis and *sai-sios* in Korean. [Niun chemkawa saisioseytayhan yenkwu.] *Eoneohak* 14. 31–51.
- Kook, Kyung-A, Ju-Won, Kim & Ho-Young Lee. 2005. A study of n-insertion preferences in Korean. [Senhoto cosalulthonghan niun chemkahyensanguy silhyenyangsang yenkwu.] *Malsori* 53. 37–60.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1984. The intersection of sex and social factors in the course of language change. Paper presented at NWAVE, Philadelphia.
- National Institute of the Korean Language [Kwuklip kwuke yenkwuwen] (ed.). 1999. *Standard Korean dictionary*. [Pyocwun kwuke taysacen.] Seoul: Doosan Dong-A. URL <http://stdweb2.korean.go.kr>.
- National Institute of the Korean Language [Kwuklip kwuke yenkwuwen] (ed.) 2014. *Sejong Text Corpus*. Retrieved from <http://www.sejong.or.kr/>
- Pagel, Mark, Quentin D. Atkinson, Andreea S. Calude & Andrew Meade. 2013. Ultraconserved words point to deep language ancestry across Eurasia. *Proc National Academy of Sciences* 110(21). 8471-8476. <https://doi.org/10.1073/pnas.1218726110>.
- Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Byber & Paul J. Hopper (eds). *Frequency and the emergence of linguistic structure*. 137–158. Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.45.08pie>.
- R Core Team. 2017. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. URL <https://www.R-project.org/>.
- Schmid, Hans Jörg (ed.). 2017. *Entrenchment, memory and automaticity: The psychology of linguistic knowledge and language learning*. Berlin: de Gruyter.
- Shin, J., & Cha, J. 2003. *Urimal Soriui Chegye (Korean Sound system)*. Sinkumunwhasa.
- Song, Jiyeon & Amanda Dalola. 2018. Crisp edge in Korean /n/-insertion. Poster session at the 7th International Conference on Phonology and Morphology. Seoul, South Korea.
- Um, Tae-su. 2010. Research of the use of n-insertion in standard Korean's regulation. *Korean Language and Literature in International Context* [kwukceyemwun] 50. 7-28.
- Vaux, B. 2002. Consonant epenthesis and the problem of unnatural phonology. In *Yale University Linguistics Colloquium*.