http://dx.doi.org/10.17959/sppm.2016.22.2.245

Contextual distribution of English loanword word-initial tensification in Korean*

Hyoju Kim
(Seoul National University)

Kim, Hyoju. 2016. Contextual distribution of English loanword word-initial tensification in Korean. *Studies in Phonetics, Phonology and Morphology* 22.2. 245-288. This study investigates the contextual distribution of word-initial tensification of English loanwords in Korean. In English loanwords, word-initial lax stops are optionally tensified. Based on two different data sources, it is confirmed that word-initial tensification of English loanwords shows three statistically significant contextual effects, which are referred to as height, length, and assimilation effect, respectively. The present study found that the word-initial tensification of English loanwords is significantly more likely to occur when the height of the vowel following the tensification site is non-high compared to high, when the word is monosyllabic rather than multisyllabic, and when the onset of the syllable following the tensification site is a tense /s'/. In order to explore the underlying reason for each contextual effect, two possible sources were investigated. The first one was the word-initial tensification of native Korean words that shows optional tensification like loanwords. The second one was the native Korean lexicon. Given the claims that the variable patterns of loanwords may reflect covert statistical generalizations of the native lexicon (Kubozono 2006, Zuraw 2010), the tendencies found in loanwords may also reflect statistical trends displayed by the Korean lexicon. The survey demonstrates that the phonological distribution of loanword tensification partially reflects the contextual distribution of both native tensification and the lexicon. In addition, acoustic and articulatory phonetic accounts are proposed with respect to the contextual effects. Finally, a learning simulation for the observed trends is conducted with a constraint set, adopting the concept of the Maximum Entropy model. The results show that the proper constraints and weights accurately predict the distribution observed in English loanwords. **(Seoul National University)**

Keywords: English loanwords, word-initial tensification, contextual distribution, Maximum Entropy model

* I would like to thank my advisor Jongho Jun and the three anonymous reviewers for their insightful comments and valuable suggestions. Without them, this manuscript would not have been developed. All remaining errors are my own.

1. Introduction

In contemporary Korean, the English voiced stops and affricate /b, d, g, ʤ/ are usually adapted as voiceless lax stops /p, t, k, c/ (e.g., [pɨripʰiŋ] 'briefing', [ticain] 'design', [kɨrup] 'group', [cənəl] 'journal'). However, for voiced stops and affricates in some English words, not only lax stops but also their tense counterparts /p', t', k', c'/ may appear in the loanwords (e.g., [panana] ~ [p'anana] 'banana', [tɛns'ɨ] ~ [t'ɛns'ɨ] 'dance', [kol] ~ [k'ol] 'goal', [cəmpʰɨ] ~ [c'əmpʰɨ] 'jump')[1].

There have been many proposals attempting to explain this variable mapping to tense stops in English loanwords. For instance, the variable mapping has been attributed to the influence of Japanese (K. Kim 1976, M. Lee 1993, Kwon 1995, among others), to the extended processes of word-initial tensification in native Korean words (Yeo 1985, Oh 2004, Shin and Davis 2004), or to the phonetic similarity between English voiced stops and Korean tense stops (Kim-Renaud 1974, S. Lee 1982, Y. Kang 2008). The majority of the previous accounts have focused mainly on the origin of the tensification in question, rather than the phonological environment or the contextual distribution of the tensification. This is probably because the word-initial tensification of English loans is widely regarded as a random idiosyncratic process that cannot be characterized by a phonological process (Yu 1987, Kwon 1995, J. Park 2008).

Only a few previous studies have discussed the phonological environment of word-initial tensification of English loanwords. Oh (2004) claimed that the tense variants may appear only when the voiced stops are word-initial and pre-vocalic in the English source words. When English voiced stops are word-medial or pre-consonantal, they are only adapted as lax stops in Korean (e.g., [təpɨl] ~ [t'əpɨl], not *[t'əp'ɨl] 'double', and [tɨrama], not *[t'ɨrama] 'drama'). In addition, Oh (2009) found that the word-initial tensification of English loans is less likely to occur before high vowels compared to non-high vowels. Oh's (2009) findings implied that the optional tensification of English loanwords is sensitive to the phonological environment of the loanwords, as opposed to their corresponding English words. However, Oh (2009) did not give much attention to loanwords themselves, but focused on phonological environment of the source language. Moreover, the author

[1] In the example words of this study, Korean tense stops are marked with ['] and the inter-sonorant voicing is not reflected.

has not provided which phonological knowledge of the speaker motivates the different tensification rates by the height of the following vowel. To this date, studies have not focused in a comprehensive investigation of the distribution of word-initial tensification with various phonological factors.

The present study has three goals. The first one is to demonstrate that the variable tensification of English loanwords in Korean, which will be called "loanword tensification," is not a random process, but rather, is a systematically governed phonological process. The second goal is to question what drives the contextual distribution of loanword tensification, focusing on two possible sources, the word-initial tensification of native Korean words, which I will call "native tensification," and the native Korean lexicon. The last goal is to examine the reason of loanword tensification based on an acoustic and an articulatory phonetic approach.

In order to accomplish the first goal, two different data sources were selected, a judgment test, and newspapers from the 1890s to 1950s. Using these two sources, this study explores whether the frequency of the tense variants, being relative to their lax counterparts, differs depending on the following three phonological factors: 1) the quality of the vowel following the target obstruent, mainly its height, 2) the word length, and 3) the laryngeal status of the syllable onset following the tensification site.

The first factor has already been discussed in Oh (2009), and this study will attempt to confirm its result. The latter two phonological factors, the word length and the onset of the following syllable, are similar to those considered in Ito's (2014) study on compound tensification in the Yanbian dialect of Korean (where tensification occurs in the onset consonant of the second element in noun compounds). In terms of word length, Ito (2014) found that the compound tensification occurs with less frequency in longer compound words, while it is more likely to occur when both elements are monosyllabic. Although the target tensification site considered in the present study is different, Ito (2014) clearly demonstrated that the word length could affect the rate of tensification. In the present study, I will examine if the word length plays a role in determining the loanword tensification rates in the word-initial position.

Similarly, the last factor was chosen from the same study conducted by Ito (2014). In her study, it is found that there is a co-occurrence restriction or OCP (Obligatory Contour Principle; Leben 1973) effect on two laryngeally marked segments observed in compound noun tensification of Yanbian Korean. More specifically, compound tensification is less likely to occur when any of the two compound members includes

a laryngeally marked segment, or an aspirated or tense consonant (e.g., /mal+tɛk'u/ [maltɛk'u] ~ [malt'ɛk'u] 'rejoinder'). On the other hand, compound tensification is more likely to occur when the members have no laryngeally marked segments (e.g., /son+ca.kuk/ [soncakuk] ~ [sonc'akuk] 'handprint').

The observed co-occurrence restriction in Korean compound tensification suggests that a similar effect may also appear in loanword tensification, although it cannot be absolute since there are loanwords with word-initial tensification and medial laryngeally marked segments (e.g., tense-tense: [pəs'ɨ] ~ [p'əs'ɨ] 'bus', [tɛns'ɨ] ~ [t'ɛns'ɨ] 'dance', [paks'ɨ] ~ [p'aks'ɨ] 'box'; tense-aspirated: [kolpʰɨ] ~ [k'olpʰi] 'golf', [təŋkʰɨ] ~ [t'əŋkʰɨ] 'dunk shot', [pəntʰɨ] ~ [p'əntʰɨ] 'bunt'). In the present study, I will consider whether the OCP effect on two laryngeally marked segments can also be found in loanword tensification.

In order to achieve the second goal, I focused on two possible sources that demonstrated the observed trends. First, in native Korean words, word-initial lax stops undergo optional tensification (e.g. /komul/ [komul] ~ [k'omul] 'junk', /cok'i/ [cok'i] ~ [c'ok'i] 'vest'). Given that voiced stops in English are usually adapted as lax stops in Korean, it might be possible that the loanword tensification is simply an extension of the process being active in native tensification. Under this possibility, the three trends found in loanword tensification might be extensions of those active processes in native tensification. Second, based on various arguments claiming that the variable patterns found in loanwords from other languages may reflect covert statistical generalizations of its native lexicon (Kubozono 2006, Luke and Lau 2008, Zuraw 2010), it is possible that the tendencies found in loanword tensification may also reflect statistical trends displayed by Korean common nouns with initial tense stops. To verify this possibility, I investigated the contextual distribution of word-initial tense stops in the Korean lexicon, focusing on the three factors which play a significant role in determining the rate of loanword tensification. Based on an in-depth discussion of the two sources, I will demonstrate that the phonological distribution of loanword tensification partially reflects the contextual distribution of both native tensification and Korean lexicon.

For the last goal of the present study, the consideration whether the phonological effects observed in loanword tensification are phonetically natural will be made. In terms of the height of the vowel following the tensification site, a relationship between the voice onset time of the preceding consonant and height of the following vowel will be discussed in-depth. Regarding the effect of the onset of the syllable

following the tensification site, I will consider Ito's (2014) explanation on the overrepresentation of the two tense onsets in simplex nouns.

The present study is organized as follows. Section 2 investigates the basic trends in English loanword word-initial tensification based on the previously mentioned data sources. It will be demonstrated that the loanword tensification is more likely to occur when initial lax stops are followed by non-high vowels, when the word in question is monosyllabic, and when the onset of the syllable following the tensification site is tense. Section 3 will examine two potential sources of the observed trends, native tensification and the Korean lexicon, to determine what drives the trends in loanword tensification. In Section 4, I will discuss whether the contextual factors observed in loanword tensification are phonetically natural, focusing on the acoustic and the articulatory phonetic motivation. Section 5 presents a learning simulation with the Maximum Entropy model (Goldwater and Johnson 2003) using the Maxent Grammar Tool (Hayes et al. 2009). The learning simulation will demonstrate that the predicted rates learned from the training data, that is English loanword data, are well matched with the observed rates. Finally, Section 6 summarizes and concludes the paper.

2. Data

To investigate the contextual factors that trigger loanword tensification, I collected data from two different sources, a judgment test and newspapers from the 1890s to 1950s. The judgment test was conducted in order to determine how speakers actually produce the initial stops of English loanwords. In addition, loanwords appearing in the newspapers are examined, since the rate of word-initial tensification of older generation loans is much higher than that of present day loans. S. Kim (2003) reported that English loans from the 1930s showed a 60% rate of word-initial tensification, present day ones are below 20%, as reported by Y. Kang (2008) and Oh (2009). Given the higher tensification rate of early-adapted English loans before 1950, earlier data is more suitable to determine the contextual effect on loanword tensification.

The investigation will be primarily concerned with three phonological environments. For each contextual effect under consideration, statistical analysis will be conducted. In particular, for the data from the judgment test, a mixed effect logistic regression model will be constructed to determine how significant each phonological factor

plays a role in determining the rate of loanword tensification.

2.1 Judgment test

I first conducted a judgment test to determine tendencies in loanword tensification. The test words used in the judgment test include 310 loanwords which are all nouns. The loanwords were selected from the loanword list compiled by the National Institute of Korean Language (NIKL 2007). The type frequencies of the selected words are provided in Table 1 below.

Table 1. Type frequencies of the word list by each subcategory

By vowel quality	/i/	/ɨ/	/u/	/ɛ/	/ə/	/o/	/a/	Total
	54	50	17	59	40	42	48	310
By word length	Monosyllabic			Multisyllabic				Total
	29			281				310
By laryngeal status	Lax		Aspirated		Tense		Sonorants	Total
	71		64		16		130	281

It is worth mentioning that the loanword list given in NIKL (2007) does not reflect the speakers' actual pronunciation as it was complied with the loanword orthography regulation enacted by NIKL in 1986, which does not allow the transcription of any English obstruent to the Korean tense stops. For this reason, loanwords given in the NIKL (2007) do not contain tense stops. Therefore, one might be concerned about the segments that are classified as a tense category in Table 1. In this study, this tense category refers to the case when the English prevocalic and word-final /s/, or the word-final /θ/ is adapted to the Korean tense /s'/. Indeed, it has already been pointed out in many studies that English prevocalic or word-final /s/ is consistently mapped to the Korean tense /s'/, e.g., [s'aicɨ] but not *[saicɨ] 'size' (Oh 1996, Y. Kang 2003, 2008). The English word-final /θ/ is also consistently borrowed into Korean as tense /s'/, e.g., [mɛs'ɨ], not *[mɛsɨ] 'math', [pɛs'ɨ], not *[pɛsɨ] 'bath' (Ahn 2003, Davis and Cho 2006).

Considering the consistent mapping of the prevocalic and word-final /s/, or word-final /θ/ in English to Korean /s'/, I classified these segments in English words as tense in loanwords, although the loans from NIKL (2007) mark those segments as lax fricative /s/ (e.g., [pəs'ɨ] 'bus', [tɛns'ɨ] 'dance', [pus'ɨ] 'booth', and [tɛs'ɨ] 'death').

On the other hand, English /s/ in a consonant cluster was classified to the lax category, because it is not realized as a tense fricative (e.g., [tɛsikʰɨ], but not *[tɛs'ikʰɨ] 'desk', [kosɨtʰɨ] but not *[kos'ɨtʰɨ] 'ghost', and [kɛsɨtʰɨ] but not *[kɛs'ɨtʰɨ] 'guest').

The test participants were 20 native Seoul-Gyeonggi Korean speakers in their twenties who have had no experience living abroad. On the survey, both the loanword form (written in Korean orthography) and its corresponding English source word (written in English) were provided for each test word. The participants were asked to choose their preferred pronunciation from the two options, lax initial and tense initial form (written in Korean orthography). They were not allowed to choose more than one option, but they could write their own response in a blank space provided on the test form in case their preferred pronunciation was not given as an option. An example of the test form is given in (1) below.

(1) Example of the judgment test form

	Loanword	English word	Option A (lax initial)	Option B (tense initial)
a.	panana	'banana'	pa.na.na	p'a.na.na
b.	cusɨ	'juice'	cu.s'ɨ	c'u.s'ɨ
c.	pusɨ	'booth'	pu.s'ɨ	p'u.s'ɨ

As shown in (1.b) and (1.c), in the case of the tense /s'/ given in option A and B, it appears as word-final /s/, and word-final /θ/ in the English source word, respectively. As I formerly mentioned, this is due to the fact that English prevocalic or word-final /s/ and word-final /θ/ are consistently adapted as the Korean tense /s'/ (Oh 1996, Ahn 2003, Y. Kang 2003, 2008, Davis and Cho 2006).

In terms of the following vowel, the judgment test results organized according to the following vowels are provided below in Table 2[2].

[2] The frequencies provided in the judgment test results are the sum of the responses for each word and for each subject.

Table 2. Tensification rate (%) in judgment test results by the following vowel

Following vowel	Lax response	Tense response	Tensification rate
i	1,054	26	2.4
ɨ	1,000	0	0.0
u	340	0	0.0
e	922	258	21.9
ə	573	207	25.9
o	693	147	17.5
a	807	173	17.7
Average	773	116	13.0

As shown in Table 2, the rate of the loanword tensification is 2.4%, 0.0%, and 0.0% before the high vowels, /i/, /ɨ/, and /u/ respectively. There is virtually no initial-tensed adaptation when the height of the following vowel is high. On the other hand, the tensification rates are 21.9%, 25.9%, 17.5%, and 17.7% before non-high vowels, /e/, /ə/, /o/ and /a/, respectively. These results are consistent with Oh's (2009) finding that word-initial tensification hardly occur when the target consonant of tensification is followed by a high vowel. The judgment test results also indicate that the word-initial tensification is blocked when the vowel following the tensification site is high, whereas it occurs much more frequently when the vowel of the initial syllable is a non-high vowel. This height effect can be seen in the following table for the tensification rates by the height of the word-initial vowels.

Table 3. Tensification rates (%) of judgment test results by following vowel height

	High	Non-high
Lax	98.9	79.2
Tense	1.1	20.8
Total response	2,420	3,780

About 1% of English loanwords with a high vowel in the initial syllable undergo tensification, whereas over 20% of loanwords with a non-high vowel show word-initial tensification. A Pearson's Chi-square test with Yate's continuity correction shows that the difference between the high and non-high vowel was statistically significant ($\chi^2 = 501.5$, df = 1, $p < 0.001$), implicating that English voiced stops and

affricate are more likely to be adapted as tense stops in Korean before non-high vowels. I will refer to this different tensification rate between high and non-high as a "height effect".

Let us now consider the distribution of the judgment test results by word length. This is shown below in Table 4.

Table 4. Tensification rates (%) of judgment test results by word length

	Monosyllabic	Multisyllabic
Lax	62.1	89.5
Tense	37.9	10.5
Total response	580	5,620

The judgment test results show that 37.9% of monosyllabic loans undergo word-initial tensification, whereas 10.5% of multisyllabic words bear a tense variant in the word-initial position. A Pearson's chi-square test with Yate's continuity correction showed that the difference between the monosyllabic and multisyllabic was statistically significant ($\chi^2 = 345.1$, df = 1, $p < 0.001$). Thus, loanword tensification is more likely to occur when the word is monosyllabic. I will refer to the difference in tensification rate depending on the word length as a "length effect".

In terms of the phonation type of the onset of the syllable following the tensification site, Table 5 shows the tensification rate and total frequencies organized by the four subcategories.

Table 5. Tensification rate (%) of judgment test results by the phonation type of the second syllable

2nd Onset / 1st Onset	Lax	Aspirated	Tense	Sonorant
Tensificaiton rate	8.7	13.8	33.8	7.0
Total response	1,420	1,280	320	2,600

As observed in the table above, the highest tensification rate was found when the second syllable began with a tense fricative (33.8%). Loanword tensification is more likely to occur when the onset of the syllable following the tensification site is a tense consonant compared to the other phonation types. In addition, the tensification rate is higher when the loans contain aspirated consonants in the onset of the second

syllable (13.8%), compared to when the second syllable consonants are lax (8.7%). The results from a chi-square independence test, given in Table 6, show that the differences among the subcategories of the second syllable onset were statistically significant, indicating that the word-initial tensification rate is correlated with the phonation type of the second syllable onset. More specifically, the differences among the obstruents were all statistically significant.

Table 6. Chi-square independence test results examining the relationship between loanword tensification and the phonation type of the second syllable onset

	χ^2	df	p-value
Among the phonation types	237.0	3	< 0.001
Among the obstruents	140.8	2	< 0.001
aspirated vs. tense	68.0	1	< 0.001
lax vs. tense	140.5	1	< 0.001
lax vs. aspirated	17.6	1	< 0.001

Based on the chi-square test results, it is confirmed that the loanword tensification occurs most frequently when the onset of the syllable following the tensification site is a tense /s'/. These results show that the tensification rate is higher when the second syllable onset is a laryngeally marked segment (i.e. tense fricative /s'/) than when it is an unmarked segment (i.e. lax consonant), suggesting that the OCP effect on laryngeally marked segments does not play a role in loanword tensification. Rather, the results imply that the co-occurrence of these two tense onsets is preferred in loanword tensification.

This apparent violation of the OCP effect on laryngeally marked segments in loanword tensification may be characterized by an assimilation type of long-distance laryngeal restrictions, which is discussed by Gallagher (2010). According to Gallagher, three types classify the laryngeal co-occurrence restrictions existing in many languages: dissimilation, assimilation, and mixed type. Among these three types, the assimilation type depicts that every obstruent given in a root must be specified for the same laryngeal feature, as is illustrated in (2) below.

(2) Assimilation type of long-distance laryngeal restrictions (Gallagher 2010): ’ stands for laryngeally marked segment, and T and K indicates obstruents with a different place of articulation.

T’-K’	*T’-K	T-K
T’-T’	*T’-T	T-T

The results of the judgment test clearly showed that loanword tensification has a preference for the co-occurrence of two laryngeally marked segments, which can be characterized by the assimilation type; T’-K’ or T’-T’. Inspired by the term used in Gallagher (2010), I will refer to this as an “assimilation effect”.

In order to evaluate how well each factor explains the observed distribution of loanword tensification, a mixed effect logistic regression model was constructed with the *glmer* function of the lme4 package (Bates et al. 2008) in R (R Development Core Team 2014). The dependent variable was each subject’s response on each test word. As the responses were either tense or lax, we can set the dependent variable as binary. Each subject and loanword were included as random intercepts. The fixed factors included the following: the height of the following vowel (high (reference), non-high), the word’s length (monosyllabic (reference), multisyllabic), and the phonation type of the onset of the syllable following the tensification site (lax (reference), aspirated, tense, and sonorant). The results are given below in Table 7.

Table 7. Results of a mixed effect logistic regression model for the judgment test: C_2 stands for the onset of the syllable following tensification site.

	Estimate	Standard Error	Pr(>|z|)
(Intercept)	-11.3324	2.2587	5.24e-07 ***
vowel height (non-high)	20.1191	3.0009	2.02e-11 ***
word length (multisyllabic)	-20.1959	2.5183	1.06e-15 ***
C_2 (aspirated)	0.4458	1.5687	0.776
C_2 (tense)	19.3574	3.6272	9.46e-08 ***
C_2 (sonorants)	-0.2894	1.5470	0.857

(Significance codes: *** ($p < 0.001$), ** ($p < 0.01$), * ($p < 0.05$) and . ($p < 0.1$))

In the first column of the table above, “vowel height (non-high)”, which indicates that following vowel height is non-high, had significantly higher ratings than the reference (i.e. vowel height (high)). The next row illustrates that monosyllabic words

(i.e. word length (mono)) showed significantly higher tensification ratings than multi syllabic words (i.e. word length (multi)). In the remaining rows, among the C_2 types, only the "C_2 (tense)" (i.e. second syllable begins with tense fricative /s'/) had significantly higher ratings than the reference (i.e. C_2 (lax)). The differences between the reference and other types of the onset, aspirated or sonorants were not statistically significant. Regarding the assimilation effect, the chi-square results showed significant differences between lax and aspirated consonants, but the results of a mixed effect logistic regression model did not show such difference. Thus, we cannot conclude that the tendency of loanword tensification consistently shows co-occurrence restrictions between the tense and aspirated consonants.

To summarize, the judgment test results have clearly shown that loanword tensification is more likely to occur when the height of the vowel following the tensification site is non-high, whereas tensification is less likely before high vowels (i.e. the height effect). When it comes to the word length, loanword tensification is more likely to occur when the word length is monosyllabic (i.e. the length effect). In terms of the laryngeal status of the onset of the syllable following the tensification site, tensification is more likely to appear when the following onset begins with the laryngeally marked segment, tense /s'/ (i.e. the assimilation effect). In the following subsection, I will continue to investigate whether these three phonological effects also play their roles actively in different data by analyzing the newspapers from the 1890s to 1950s.

2.2 A survey of previous newspapers

In this section, I will continue to examine more trends in loanword tensification from another data source, newspapers that were published between the 1890s and 1950s. A large number of English words were borrowed in this early adaptation stage (Ito et al. 2006, N. Kim 2012). During this period, initial voiced stops or affricate in the English source words were adapted as Korean tense stops in many cases, when such initial tense stop adaptation does not appear in loanwords adapted in the present day. The loanword tensification rate was much higher than that of present day. This is demonstrated by the tensification rate of early loans being about 60% (S. Kim 2003), whereas now are much lower as was seen in the judgment test results (11.4%) (cf. Y. Kang (2008) reported that around 20% of loanwords undergo word-initial tensification). The examples given in (3) demonstrate this frequent adaptation of

English voiced stops and affricates to Korean tense stops as seen in Korean newspapers from the1890s to 1950s.

(3) English loans taken from newspapers printed during the 1890s to 1950s (the newspaper name and the issued date are in parenthesis)

a. [k'oltɨ]	'gold'	(*Dongailbo* 07-12-1937)
b. [t'ainamait^{h}ɨ]	'dynamite'	(*Chosunilbo* 03-05-1931)
c. [t'əkillasɨ]	'Douglas'	(*Dongailbo* 26-08-1938)
d. [p'alomɛt^{h}ə]	'barometer'	(*Dongailbo* 16-12-1934)
e. [c'ənalicɨm]	'journalism'	(*Chosunilbo* 01-07-1938)

The early-adapted loanwords in (3) generally have corresponding words with no initial tensification in present day Korean. For instance, all participants in the judgment test failed to adapt 'gold' with the initial tense variant, but rather adapted it only with a lax initial variant (e.g., [koltɨ] but not *[k'oltɨ] 'gold'). Since there were more words undergoing tensification than the data from my own judgment test, there were a much greater number of phonological environments that affected loanword tensification. Hence, it would be much more efficient to examine which phonological factors have an influence on the loanword tensification from this older data.

To investigate the tendencies in loanword tensification of the earlier loans, I now introduce a list of loans printed in newspapers from the 1890s to 1950s, which was composed by A. Kim (2009). According to A. Kim (2009), the words were collected from five different newspapers: *Doklipshinmun, Daehanmaeilshinbo, Hwangsungshinmun, Chosunilbo,* and *Dongailbo.* The word list does not contain all the printed loanwords. In the case of *Chosunilbo* and *Dongailbo,* data before the 1920s is unavailable since they began printing after 1920. Additionally, in the case of *Chosunilbo* and *Dongailbo*, twelve newspapers a year were randomly selected due to the large amount of relevant data[3]. Of all the 3,623 tokens of A. Kim (2009), 458 tokens (type frequency: 142) having an initial voiced stops and affricate in English, were analyzed to examine whether the three phonological effects, the place, height, length, and

[3] A. Kim (2009) followed three criteria for loanword selection. First, loanwords were chosen only when both pronunciations and spellings of their corresponding English words were known. Second, names of people and places that had very low frequencies were excluded. Finally, loanwords that are clearly regarded as Japanese origin loanwords were excluded.

assimilation effects, also actively affected actively the distributions of loanword tensification in the early-adapted loanwords.

In terms of the height effect, the distribution of tensification organized by height of the vowel following the tensification site is provided in Table 8.

Table 8. Loanword tensification rate (%) of newspapers from the 1890s to 1950s by following vowel height

	High	Non-high
Lax	79.4	53.2
Tense	20.6	46.8
Total frequency	131	327

As we can see in the table above, while the tensification rate is 20.6% when the tensification site is followed by a high vowel, the rate is 46.8% when the word-initial stop is followed by a non-high vowel. A Pearson's Chi-square test with Yates' continuity correction indicated that this difference is statistically significant (χ^2 = 25.8, df = 1, p = 0.001). Therefore, the height effect is also confirmed as a significant factor in the newspapers data.

Table 9. Loanword tensification rate (%) of newspapers from the 1890s to 1950s by syllable count:

	Monosyllabic	Multisyllabic
Lax	23.8	62.5
Tense	76.2	37.5
Total frequency	21	437

Turning now to word length (length effect), Table 9 above shows a higher loanword tensification rate in monosyllabic words (76.2%) than that of multisyllabic words (37.5%). A Pearson's chi-square test with Yate's continuity correction showed the difference between monosyllabic and multisyllabic tensification to be statistically significant (χ^2 = 10.9, df = 1, p < 0.001), providing further proof that the length effect plays a significant role in the loanword tensification[4].

[4] Considering the small sample size of monosyllabic words (token frequency of monosyllabic-lax category was 5), Fisher's exact test was conducted as well. The significant difference

Finally, in relation to the assimilation effect, English loanwords appearing in newspapers from the 1890s to 1950s again showed the highest tensification rate when the onset of the syllable following the tensification site was the tense fricative /s'/.

Table 10. Loanword tensification rate (%) of newspapers from the 1890s to the 1950s by second syllable onset

2nd Onset / 1st Onset	Lax	Aspirated	Tense	Sonorants
Lax	61.2	74.5	31.1	70.3
Tense	38.8	25.5	68.9	29.7
Total frequency	80	47	74	236

The differences between the second syllable subcategories were statistically significant, suggesting that the tensification rate of the early loans was influenced by the onset of the syllable following the tensification site. Specifically, the results showed that the tensification rate was significantly higher when the onset of the following syllable was tense (obstruents: $\chi^2 = 24.8$, df = 2, $p < 0.001$; aspirated vs. tense: $\chi^2 = 19.1$, df = 1, $p < 0.001$; lax vs. tense: $\chi^2 = 13.3$, df = 1, $p < 0.001$), and the tensification rate difference between lax and aspirated consonants was not significant (lax vs. aspirated: $\chi^2 = 1.43$, df = 1, $p = 0.2$). Therefore, we can conclude that the assimilation effect also plays a determining role in tensification, as is supported by the newspaper data from the 1890s to 1950s.

Table 11. Chi-square independence test results on the relationship between loanword tensification observed in the newspapers and the onset of the syllable following the tensification site

	χ^2	df	p-value
Among the phonation types	40.2684	3	< 0.001
Among the obstruents	25.1644	2	< 0.001
aspirated vs. tense	19.9767	1	< 0.001
lax vs. tense	12.8688	1	< 0.001
lax vs. aspirated	1.7573	1	0.185

between monosyllabic and multisyllabic remains the same (The Fisher's exact test statistic value is 0.0008; the results ars significant at $p < 0.01$).

The results from the newspaper data clearly demonstrate that the height, length and assimilation effects all play determining roles in loanword tensification. These results suggest that the three effects, which have been proven to play significant roles in the present day loanword tensification, also applied actively in the tensification of older adapted loanwords. In other words, although the rates of present day loanword tensification have greatly decreased compared to early adaptations, the three phonological factors have been and continue to affect loanword tensification.

To summarize this section, three contextual factors under consideration (i.e. height, length, and assimilation effects), were supported by the results from the two different data sets. Table 12 summarizes the trends found in loanword tensification according to the two data sources.

Table 12. Summary of findings

factors / Data	Following vowel height	Word length	Second syllable onset
Judgment test	non-high > high	monosyllabic > multisyllabic	tense > aspirated > lax > sonorants
Newspaper survey	non-high > high	monosyllabic > multisyllabic	tense > aspirated / lax > sonorants

The effects of the three phonological factors on loanword tensification were found to be statistically significant in all data. Based on the observations from two different data sets, the tendencies uncovered in the distribution of loanword tensification are summarized as follows.

(4) Trends in loanword tensification

a. Loanword tensification is more likely to occur when the vowel following the tensification site is a non-high vowel /ɛ, ə, o, a/, and it is less likely to appear before high vowels /i, ɨ, u/. (height effect)
b. Loanword tensification is more likely to occur within monosyllabic words than within multisyllabic words. (length effect)
c. Loanword tensification is more likely to occur when the onset of the syllable following the tensification site is tense /s'/. (assimilation effect)

Now we can question where these tendencies come from. Why are word-initial

voiced stops or affricate in English more likely to undergo tensification in loanword adaption in Korean when the stops are followed by non-high vowels, when the words are monosyllabic, and when the onset of the syllable following the tensification site is tense? In the following section, I will discuss two possible sources of the observed trends. First, in Korean, word-initial lax stops undergo optional tensification (e.g., /komul/ [komul] ~ [k'omul] 'junk', and /cok'i/ [cok'i] ~ [c'ok'i] 'vest'). Given that voiced stops in English are usually adapted as lax stops in Korean, it might be possible that loanword tensification is simply an extension of the tensification process active in native Korean words, which I referred to as "native tensification" in section 1. In that case, the three trends found above might be extensions of those active in native tensification. In order to explore this possibility, I investigated the distribution of native tensification.

Second, it is possible that the trends in loanword tensification may reflect the statistical trends displayed by native words with initial tense obstruent in the Korean lexicon. To verify this hypothesis, I will examine the contextual distribution of word-initial tense stops in the Korean lexicon, focusing on the three factors which play a significant role in determining the rate of tensification in English loanwords.

3. What drives the trends in loanword tensification?

In this section, I will explore the causes of the three contextual effects found in loanword tensification, focusing on distributions existing in two different sources; native tensification, and native Korean lexicon. Loanword and native tensification share the target and output consonant types, and even the position of the target. If there were such an effect in native tensification, this would suggest that the phonological effects found in loanword tensification are simply generalized from the same phonological effects active in native tensification. With this in mind, I will first investigate the contextual distribution occurring in native tensification, and then compare its results with those of loanword tensification. If these two tensification processes show similar tendencies, it is possible to say that the trends in loanword tensification are influenced by the phonological distributions of native tensification.

The second possible source that may drive the trends in loanword tensification is the distribution of tense consonants in Korean simplex words. The effect of the native lexicon on the emergent patterns in loanwords has been discussed often in the literature (Herd 2005, Rose and Demuth 2006, Y. Kang 2011 among others). More

specifically, the variable patterns in loanwords are affected by the covert statistical generalizations of the native lexical distribution (Kubozono 2006, Luke and Lau 2008, Zuraw 2010). For instance, Zuraw (2010) showed that the nasal substitution of Spanish loanwords in Tagalog directly reflects the statistical trends in the native Tagalog lexicon. In addition, Luke and Lau (2008) found a monosyllabicity preference for verbs and bisyllabicity preference for nouns in the truncation of recent English loans in Cantonese. They demonstrated that this noun–verb asymmetry is consistent with the lexical statistics in native Cantonese words.

In this regard, it is possible that the tendencies in loanword tensification may reflect the statistical trends displayed by native words with initial tense stops in the Korean lexicon. If it is true that the trends in loanword tensification are generalized from the extant distribution existing in Korean nouns, the distribution of Korean nouns has to be characterized by each effect as follows: 1) the height effect reflects the underrepresentation of word-initial tense stop being followed by high vowels in Korean common nouns, 2) the length effect mirrors the overrepresentation of word-initial tense stops in monosyllabic words of Korean common nouns, and 3) the assimilation effect reflects the overrepresentation of word-initial tense stops when the onset of the following syllable is a tense consonant. In the following subsections, I will investigate distributions of the two possible sources in order to examine which hypothesis is more suitable for explaining the tendencies in loanword tensification.

3.1 Does native tensification show the same trends?

In this section, I will explore whether the three phonological factors found in loanword tensification actively play a role in native tensification. Before investigating the trends in native tensification, I will briefly discuss previous studies on native tensification and previous accounts on the relationship between loanword tensification and native tensification.

Some Korean words undergo optional word-initial tensification in a similar way to loanword tensification as shown in (5).

(5) Examples of native tensification

a. /komul/	[komul] ~ [k'omul]	'junk'
b. /cat^{h}uli/	[cat^{h}ɯri] ~ [c'at^{h}ɯri]	'remnant'
c. /cok'i/	[cok'i] ~ [c'ok'i]	'vest'

d. /kockam/ [kotk'am] ~ [k'otk'am] 'dried persimmon'
e. /toŋkiɬami/ [toŋkiɾami] ~ [t'oŋkiɾami] 'circle'

Word-initial tensification in native Korean words is a minor phonological process. According to D. Park (2000), 233 out of 22,166 existing words (about 1.0%), show word-initial tensification. Many proposals have been made on this optional tensification in native words (Heo 1965, S. Lee 1981, M. Lee 1989, D. Park 2000, S. Kim 2001, Han 2010). However, the majority of the previous accounts have considered native tensification as a random phonological process, concluding that it cannot be defined by a phonological rule. Rather, it has been proposed that native tensification is a process of semantic variation, such as adding a negative nuance (e.g., [colpjəŋ] 'soldier' ~ [c'olpjəŋ] 'servant'), or adding an emphatic nuance (e.g., [cinhan] 'thick' ~ [c'inhan] 'thicker') (D. Park 2000).

Furthermore, nearly all-relevant studies did not pay much attention to the relationship between native tensification and loanword tensification, since the latter is usually regarded as a simple extension of the former (Yeo 1985, Oh 2004, 2009, Shin and Davis 2004). For a more detailed comparison between loanword tensification and native tensification, I examined whether the phonological factors that play significant roles in loanword tensification also actively affect native tensification. If there are such effects in native tensification, the phonological trends found in loanword tensification are simply generalized from the same phonological trends active in optional tensification appearing in Korean words.

3.1.1 Survey

For the purpose of investigating the contextual distribution of native tensification, I conducted a judgment test with 210 native Korean words, which were reported in previous studies to undergo native tensification in the Seoul-Gyeonggi dialect. These test words were selected from word lists with native tensification in D. Park (2000) and Han (2010). They include nouns, verbs, adjectives, adverbs, and a determiner as shown in Table 13.

Table 13. Classification of testing words by part of speech

Word class	Noun	Verb	Adjective	Adverb	Determiner	Total
Number of words	106	54	34	15	1	210

These 210 words were tested by twenty participants who were native Seoul Korean speakers in their twenties, with no experience living abroad. All participants had lived in the Seoul-Gyeonggi area for more than twenty years. In the judgment test, both initial lax and initial tense forms for each test word were provided as options (written in Korean orthography). The participants were instructed to choose the pronunciation form used in their everyday life between two possible forms, and multiple choices were not allowed. If their pronunciation did not correspond to one of the options, they were asked to write their own pronunciation in a blank space of the survey form. An example of the survey form is provided in (6).

(6) Example of the survey form

	Word		Option A	Option B
a.	cakta	'be small'	cakt'a	c'akt'a
b.	kockam	'dried persimmon'	kotk'am	k'otk'am
c.	toŋkiɬami	'circle'	toŋkiɬami	t'oŋkiɬami

It should be noted that the pronunciations given in the options (except the target tensification site) follows the standard pronunciation provided in the Standard Korean Dictionary (henceforth, SKD)[5]. The pronunciations given in SKD reflect a phonological process, such as the so-called Korean Post Obstruent Tensing rule, which illustrates lax obstruents of the onset following an obstruent coda as tense. Reflecting this characteristic of SKD, the second syllable initial lax stops of the words given in (6.a) and (6.b) are presented as tense stops in the options (e.g., /cak.ta/ [cak.t'a] 'be small' and /koc.kam/ [kot.k'am] 'dried persimmon').

[5] SKD is available at http://stdweb2.korean.go.kr/search/List_dic.jsp.

3.1.2 Results

The survey results on native tensification were analyzed for each contextual effect under consideration (i.e. the height, the length, and the assimilation effects). I first examined whether the height effect also played a role in native tensification. The results by the height of the vowel following the tensification site are presented in Table 14.

Table 14. Native tensification rate (%) by following vowel height

	High	Non-high
Lax	75.3	62.4
Tense	24.7	37.6
Total response	1,240	2,960

The rate of native tensification was 24.7% when the following vowel was high, whereas the native tensification occurred at a rate of 37.6% when the height of the following vowel was non-high. A Pearson's chi-square test with Yates' continuity correction showed that the difference between high and non-high tensification was statistically significant ($\chi^2 = 64.4$, df = 1, p-value < 0.0001), indicating that the height effect was also confirmed as significant in the native tensification.

Nevertheless, when considering the number of syllables, tensification rates remain almost the same regardless of the syllable count. Indeed, the tensification rates actually seem to increase slightly in words with multiple syllables.

Table 15. Native tensification rate (%) by word length

	Monosyllabic	Multisyllabic
Lax	67.9	66.2
Tense	32.1	33.8
Total response	140	4,060

A Pearson's chi-square test with Yate's continuity correction indicated that the difference between the monosyllabic and multisyllabic tensification was not statistically significant ($\chi^2 = 0.1031$, df = 1, p = 0.7481). Thus, the rate of native tensification does not appear to be significantly related to word length, which means that the length effect is not confirmed in native tensification.

Next, I inspected the tensification rate by the onset of the syllable following the tensification site. The results showed that, unlike the loanword tensification data, native tensification rate with tense consonants in the second syllable onset was not the highest. The following data in Table 16 shows that the highest rate of tensification occurs in the aspirated category (35.3%) rather than in the tense category (34.6%).

Table 16. Native tensification rate (%) by second syllable onset type

2nd Onset / 1st Onset	Lax	Aspirated	Tense	Sonorant
Tensification	34.5	35.3	34.6	31.3
Total response	1,600	320	1,160	980

The differences between the subcategories of the second syllable onset were not statistically significant, suggesting that there was no association between the native tensification and the onset of the second syllable. More specifically, none of the subcategories among the obstruents showed significant differences. Therefore, the assimilation effect does not appear to have any bearing on native tensification.

Table 17. Chi-square independence test results on the relationship between native tensification and the onset of the syllable following the tensification site

	χ^2	df	p-value
lax *vs.* aspirated *vs.* tense *vs.* sonorant	3.66	3	0.3003
lax *vs.* aspirated *vs.* tense	0.07	2	0.9612
aspirated *vs.* tense	0.03	1	0.85
lax *vs.* tense	0	1	1
lax *vs.* aspirated	0.04	1	0.83

For the native tensification data, a mixed effect logistic regression model analysis was performed via R (*glmer* function in lmer4 package of the R statistical software, R Development Core Team 2014) in order to examine how well each factor could explain a native tensification model. Subjects and test words were included as random effects. The fixed factors are the same as the factors adopted in the model for the judgment test on loanword tensification.

Table 18. Results of a mixed effect logistic regression model for native tensification: C_2 stands for the onset of the syllable following the tensification site.

	Estimate	Standard Error	Pr(>|z|)
(Intercept)	-2.3010	1.0428	0.0273 *
vowel height (non-high)	1.4618	0.4175	0.0004 ***
word length (multi)	0.0260	1.0497	0.9801
C_2 (aspirated)	-0.2855	0.7322	0.6966
C_2 (tense)	-0.2001	0.4587	0.6625
C_2 (sonorants)	-0.6754	0.4885	0.1668

(Significance codes: *** ($p < 0.001$), ** ($p < 0.01$), * ($p < 0.05$) and . ($p < 0.1$))

As expected and revealed by the chi-square test analysis, the mixed logistic regression model shows that the only significant factor is the vowel's height following the tensification site, indicating that only the height effect is confirmed as bearing significant influence on both native tensification and loanword tensification. On the other hand, other effects (i.e. the length and assimilation effect) do not seem to bear any significant influence on the native tensification model.

3.1.3 Summary

In summary, the tendencies in native tensification are partially attributed to the height of the following vowel, which implies that the height effect holds for both native tensification and loanword tensification. However, unlike the loanword tensification, the native tensification did not demonstrate any significant length effect nor assimilation effect. The trends in native tensification are summarized in Table 19.

Table 19. Trends in native tensification

Factors / Data	Following vowel height	Word length	Second syllable onset
Judgment test	non-high > high	Not applicable	Not applicable

In other words, it seems that the contextual distribution of native tensification is generally different from that of English loanwords. Although it is possible to argue that loanword tensification actively reflects the same height effect observed in the

native phonological process (i.e. native tensification), we can still question the reason for the significant length effect and assimilation effect. In the following section, the question whether the trends in loanword tensification are affected by the distribution of the native word lexicon will be explored.

3.2 Do the trends in loanword tensification reflect the distribution of the Korean lexicon?

In this section, I continue to investigate what might affect the trends observed in loanword tensification, focusing on the contextual distribution of word-initial stops in the Korean lexicon. A potential explanation of the tendencies in loanword tensification can be found in the distribution of native words. So, if the three phonological factors influencing the loanword tensification are reflected in the distribution of native words, it is possible to contend that the trends in English loanwords reflect the existing contextual distribution of the native word lexicon.

3.2.1 Survey

In order to investigate the distribution of the Korean lexicon by the contextual factors found in loanword tensification, Korean common nouns were collected from the *Frequency of Korean* (B. Kang and H. Kim 2009), which contains 117,333 type frequencies, and 7,889,661 token frequencies. From the corpus of B. Kang and H. Kim (2009), 34,501 common nouns registered in SKD with a frequency value greater than five were selected. Among these 34,501 common nouns, 13,252 words (including 12,843 words with initial lax stops /p, t, k, c/ and 409 words with initial tense stops /p', t', k', c'/) were used to examine the distribution of native words. For the selected data, the ratio of observed frequency to expected frequency (henceforth, O/E) was calculated by each category. If the O/E values are greater than 1.0, this indicates that the category occurs more than expected (i.e. overrepresentation), whereas the O/E values less than 1.0 indicate that the category occurs less than expected (i.e. underrepresentation). The detailed calculation process for the O/E values will be provided in following subsection.

3.2.2. Results

Based on the selected word list, I calculated the ratio of observed frequency to expected frequency in order to examine whether the distribution of native words is sensitive to each factor found in loanword tensification. I first observed the distribution of native words by the place of articulation of word-initial stops. Prior to examining the distribution of native words by a following vowel, it should be remembered that loanword tensification is less likely to occur when the height of the vowel following the tensification site is high. If the distribution of Korean common noun shows an underrepresentation of the initial tense stop being followed by high vowel, the height effect of loanword tensification could be generalized from the distribution of native word lexicon. Table 20 shows the observed number of Korean noun organized by the vowel height following the initial stop

Table 20. Observed number of Korean nouns by the vowel height following the initial stop

	High	Non-high	Total
Lax	4,289	8,554	12,843
Tense	93	316	409
Total	4,382	8,870	13,252

Based on the observed frequencies in Table 20, the expected frequency of each cell was calculated. The expected number for each (x, y) pair is written as E(x, y) = P(x) ×P(y)×N, where P(x) is the probability of (x), P(y) is the probability of (y), and N is the total number of each cell (Kawahara et al. 2006). For example, looking at the Table 20, the probability of tense stops occurring in word-initial position is 409/13,252 = 0.0308, and the probability of a high vowel occurring as the vowel following a word-initial stop is 4,382/13,252 = 0.3306. This being the case, the probability of tense stops being followed by high vowel is 0.0308×0.3306 = 0.0102. Thus, the expected frequency of tense stops followed by a high vowel is 0.0102 ×13,252 = 135.2. The expected number and the O/E values for common Korean nouns organized by the following vowel height are provided in Table 21.

Table 21. Expected numbers and O/E values for Korean nouns by the following vowel height: O stands for observed frequency and E for expected frequency

	High	Non-high	Total
Lax	E = 4246.8 O/E = 1.01	E = 8596.2 O/E = 1.00	12,843
Tense	E = 135.2 O/E = 0.69	E = 273.8 O/E = 1.15	409
Total	4,382	8,870	13,252

As we can see in the shaded box in Table 21, the O/E values for initial tense stops with a following high vowel are lower than 1.0, whereas tense stops with a following non-high vowel are higher than 1.0. These results indicate that the word-initial tense stops being followed by a high vowel are underrepresented, while those being followed by a non-high vowel are overrepresented. In addition, a chi-square independence test showed that the difference between high and non-high is significant (χ^2 = 19.8, df = 1, $p < 0.001$), suggesting that the distribution of word-initial lax and tense stop is correlated with the following vowel's height.

The underrepresentation of tense stops being followed by high vowel further supports the height effect of the loanword tensification. That is, a possible reason for the low tensification rate before a high-vowel in English loanwords may be simply because the tense consonants are less likely to be followed by a high vowel in the native Korean lexicon. Thus, the underrepresented pattern of the initial tense consonants with following high vowels in comparison to the overrepresented pattern of the initial tense consonants being followed by non-high vowels could be one reason for the lower rate of loanword tensification when the vowel following the tensification site is high. This implies that the variable tensification patterns in the English loanword could be generalized from the existing lexical distribution of the common nouns in Korean.

Although we cannot emphatically conclude that the height effect in the loanword tensification is fully generalized by the native lexical distribution, it seems that the lexical distribution of the Korean common nouns might be one answer to the question of what drives the significant role of the height effect in loanword tensification.

Another significant phonological factor found in loanword tensification was the length effect. English loanwords showed a significantly higher tendency of word-

initial tensification when the word was monosyllabic compared to multisyllabic. One possible reason for this might be the distribution in the existing Korean lexicon. If monosyllabic words starting with tense stops are overrepresented in the lexicon, it is not too much to say that the native word distribution has an influence on the length effect of the loanword tensification. In order to verify this hypothesis, O/E values by word length were calculated based on the same word list of Korean common nouns from B. Kang and H. Kim (2009). The results are illustrated in Table 22.

Table 22. O/E values for common Korean nouns by syllable count

	Monosyllabic	Multisyllabic
Lax	O = 330 O/E = 0.91	O = 12,513 O/E = 1.00
Tense	O = 46 O/E = 3.96	O = 363 O/E = 0.91

The difference between word-initial tensification in monosyllabic and multisyllabic words was statistically significant ($\chi^2 = 105.1$, df = 1, $p < 0.001$), suggesting that the distribution of word-initial lax and tense stops is dependent on the word length. The O/E value of the monosyllabic-tense category (shaded box in Table 22) indicates that the word-initial tense stops in monosyllabic words are highly overrepresented, whereas the multisyllabic-tense category is underrepresented (O/E = 0.91). The high O/E value of the monosyllabic-tense category suggests that the word-initial tense stops in monosyllabic words are highly overrepresented in Korean common nouns, indicating that the length effect of loanword tensification may reflect the distribution observed in native Korean words.

As presented in section 2, the differences among the subcategories of the onset of the syllable following the tensification site were in their effects on loanword tensification. More specifically, loanword tensification is more likely to occur when the onset of the syllable following the tensification site is tense /s'/, implying that the observed tendencies seem to be regarded as an assimilation type of long-distance laryngeal co-occurrence restriction. What, then, drives the assimilation effect observed in loanword tensification? As a possible hypothesis, the assimilation effect might be attributed to the frequent co-occurrence of two tense pairs in native lexical entries. If this hypothesis is correct, then we may conclude that the assimilation effect of loanword tensification is generalized from the distribution which exists across the

native word lexicon. For the analysis of the lexical distributions of Korean common nouns by the onset of the syllable following the word-initial stops, the same word list was used. It should be noted that the monosyllabic words are excluded because they lack of a second syllable. The O/E values for onset pairs are provided in Table 23.

Table 23. O/E values for onset pairs

2nd Onset / 1st Onset	Lax	Aspirated	Tense	Sonorants
Lax	O = 6,341 O/E = 1.02	O = 1,356 O/E = 1.01	O = 1,442 O/E = 0.92	O = 3,374 O/E = 1.00
Tense	O = 74 O/E = 0.41	O = 32 O/E = 0.82	O = 169 O/E = 3.72	O =88 O/E = 0.90

The O/E values of the second syllable tense consonants are the highest (O/E = 3.72) among the words with word-initial tense stops. This suggests that the co-occurrence of the tense-tense onset pair is highly overrepresented. Given the observed frequencies, the differences among the subcategories of the second syllable onset were statistically significant (χ^2 = 413.3, df = 3, $p < 0.01$), meaning that the type of the first onset and the second onset consonants are correlated in Korean common nouns. More specifically, the differences among the obstruents and the differences between tense and other onsets were all statistically significant (among the obstruents: χ^2 = 398.0, df = 2, $p < 0.01$; lax vs. aspirated: χ^2 = 10.4, df = 1, $p < 0.01$; lax vs. tense: χ^2 = 379.1, df = 1, $p < 0.01$; aspirated vs. tense: χ^2 = 78.5, df = 1, $p < 0.01$), which reveals that the phonation type of the second syllable onset and the word-initial stops are correlated.

It should be mentioned that this observation in Korean common nouns is consistent with other studies including H. Kang and Oh (2015), and Ito (2014). In H. Kang and Oh (2015), they investigated 67,258 words from the Standard Korean Pronouncing Dictionary, which was constructed by H. Lee (2002). Based on the observed frequency of the onsets in the first and second syllables, the O/E values were calculated. The results of H. Kang and Oh (2015) are provided in Table 24.

Table 24. O/E values for onset pairs in the first and the second onsets (H. Kang and Oh, 2015)

2nd Onset / 1st Onset	Lax	Tense	Aspirated	Nasal	Liquid	Vocoid
Lax	1.03	0.91	1.01	1.00	0.96	1.05
Tense	1.04	2.35	0.33	0.75	1.38	0.74
Aspirated	1.00	0.95	1.07	1.02	0.86	0.98

Excluding the ratios of tense-tense pairs and tense-aspirated pairs, the O/E ratios in Table 24 are all approximately 1.00. The ratio of tense pairs bears the highest ratio. My observation on Korean common nouns also confirms the highest O/E value for the tense-tense category (O/E = 3.72). Given this, the tendencies found in loanword tensification may be generalized from the overrepresented tense pairs in native words.

Further supporting evidence for the overrepresentation of the co-occurrence of two tense onsets in the lexical distribution of Korean nouns can be found in Ito (2014). In her study of Korean compound nouns, she investigated the O/E values for onset types in initial and peninitial syllables in order to see if the OCP effect appeared in compound tensification of Korean. Table 25 shows the results of Ito (2014).

Table 25. O/E values for onset types in initial and peninitial syllables in disyllabic and trisyllabic simplex nouns; significant correlations ($p < 0.05$) are indicated by * (Ito 2014).

2nd Onset / 1st Onset	Aspirated	Tense	Lax	Sonorant	Total
Aspirated	0.44 *	0.80	1.09	1.11	141
Tense	0.94	2.30 *	0.83	0.80	267
Lax	1.08	0.82 *	1.00	1.04	1624
Sonorant	0.97	0.97	1.05	0.96	843
Total	275	333	1236	1031	2875

The finding of Ito (2014) also denotes that the lexical distribution of Korean nouns shows a significant overrepresentation of tense-tense onset pairs in initial and following syllables. These findings have clearly shown that the overrepresentation of the tense-tense onset pairs in initial and following syllables is a salient characteristic of Korean nouns.

The assimilation effect of loanword tensification has illustrated that the word-

initial lax stops of loanwords vary with tense stops. This effect drives loanwords to have one more laryngeally marked segment, even though the word already contains a laryngeally marked segment in the second syllable onset. Given the observations from the present study and from the previous studies, this tendency in loanword tensification may reflect a lexical trend in Korean nouns, which shows an overrepresentation of two tense onset pairs despite having more laryngeally marked segments within a word. That is, the assimilation effect of the loanword tensification has been generalized from the existing lexical trends in Korean words.

3.2.3 Summary

In this section, I have examined whether the distribution of the native Korean lexicon affects the trends observed in loanword tensification. The distribution of Korean common nouns can be characterized as follows.

(7) Trends in Korean common nouns
a. Underrepresentation of word-initial tense stops being followed by high vowels
b. Overrepresentation of word-initial tense stops in monosyllabic words
c. Overrepresentation of two tense obstruents in the first and second syllable onsets

This lexical distribution of Korean common nouns suggests that the loanword tensification is partially affected by the trends in the native Korean lexicon. In particular, the significantly low rate of loanword tensification before high vowels (i.e. height effect) might be influenced by the underrepresentation of word-initial tense stops being followed by high vowels in Korean common nouns. In terms of the word length, the high rate of loanword tensification in monosyllabic words (i.e. length effect) may be affected by the overrepresentation of word-initial tense stops in monosyllabic words. Finally, the high rate of loanword tensification when the onset of the syllable following the tensification site is tense /s'/ (i.e. assimilation effect), is possibly affected by a high overrepresentation of two tense onset pairs in Korean nouns.

Based on the findings in the native word lexicon, it is likely that the contextual distribution of loanword tensification partially reflects salient patterns existing in the distribution of native word lexicon. Then, we should ask ourselves whether the three phonological factors found to influence loanword tensification are phonetically

natural. In the following section, I will discuss the phonetic naturalness of the effects with supporting evidence.

4. Discussion

Based on two different data sources, I have thus far demonstrated that three phonological factors - height effect, length effect, and assimilation effect - have a significant influence on the distribution of loanword tensification. We have also observed a similar process of variation appearing in native tensification while exploring the question of what drives the contextual distribution of loanword tensification. From the survey results, the distribution of native tensification has partially shown the same effects in loanword tensification; the place and height effect. Then, the distribution of Korean common nouns was investigated as another possible source of loanword tensification. The covert distribution of Korean common nouns has shown that all the effects found in loanword tensification are generalized from the overrepresented distribution in the native word lexicon.

In this section, I will keep discussing possible motivations for each factor observed in the distribution of loanword tensification. Specifically, the main goal of this section is to consider the phonetic naturalness of the contextual effects, focusing on the acoustic and articulatory motivations of the observed phonological factors.

4.1 The height effect

As we have seen so far, the height effect has been a significant factor of both loanword tensification and native tensification. Furthermore, the examination on Korean common nouns has indicated that this height effect may reflect the extant patterns of the native word lexicon; underrepresentation of word-initial tense stops being followed by high vowels. To what, then, might we attribute the effect of the following vowel's height on the word-initial stop? Specifically, is the height effect phonetically natural?

One response to this question can be drawn from acoustic motives, especially from the relation between the voice onset time (VOT) of the preceding stop and the height of the following vowel. Since the height effect of the loanword tensification has illustrated a variation of preceding consonants in accordance with the height of the following vowel, we can reason that there is a vowel-dependent variation in the

consonant.

In terms of the VOT duration of Korean stop series, tense stops are shorter (6~18ms, Silva 2006) than lax (36~90ms, Silva 2006). On the basis of this characteristic, if the VOT duration increases as the height of the following vowel increases, it would suggest that higher vowels might avoid tense stops that have the shortest VOT. On the other hand, if the higher height of the following vowel forces the preceding stop to have a shorter VOT value, this would imply that the higher vowel prefers the shorter VOT duration for the preceding stop. This hypothesis with respect to the phonetic naturalness of the height effect can be examined through several previous studies on the relation between the VOT value of word-initial stops and the height of the following vowel.

Among previous studies, Klatt (1975) has found that the VOT value of the preceding stop is slightly longer before high vowels, and shorter before low vowels. As for this tendency between the VOT parameter and the height of vowel, Klatt (1975) has claimed that high vowels influence the larynx, resulting in a difficult voicing. This influence can be attributed to the fact that the raising of the tongue body can result in a slight upward pull on the larynx, which in turn may be harder for phonation to start. Thus, a slightly higher subglottal pressure would be required to initiate phonation.

In addition, Ohala (1981) has also found that VOT becomes longer before high vowels and shorter before mid or low vowels. Ohala (1976, 1981) has suggested that the effect of the following vowel's height occurs when high vowels offer greater resistance to the air escaping from the oral cavity, which delays the achievement of a trans-glottal pressure suitable for voicing. Likewise, Docherty (1992) has supported this relationship between the VOT duration of the preceding stop and the height of the following vowel. His experiment showed that the vowel quality is a significant factor with respect to the VOT parameter since a longer VOT duration of both English voiced and voiceless stops occurs more before high vowels compared to low vowels[6].

---

[6] It should be noted that the account on the relationship between VOT and the following vowel has already been attempted by Oh (2009). The author has pointed out that the trends of the initial stops in VOT and the height of the following vowels can be applied to tense vowels but not to lax vowels because the high lax vowels are associated with the shortest VOT. For this reason, English high lax and tense vowels behave the same way in loanword tensification since they rarely trigger loanword tensification even though lax high vowels

These previous studies provide a firm foundation to support the observed height effect found in loanword tensification. Due to the necessarily shorter VOT for tense consonants, the longer VOT before high vowels implies that the occurrence of the tense variant being followed by high vowels is discouraged. In contrast, the shorter VOT before non-high vowels suggests that the preceding lax stops are prone to be realized as tense variants when the height of the following vowel is non-high. Thus, the previous studies offer further support for the height effect found in loanword tensification and show that the height effect is indeed phonetically natural. However, it is necessary to confirm that these characteristics between VOT and the following vowel is also applied to Korean stops and following vowel.

4.2 The assimilation effect

As it has been demonstrated in section 2, the rate of the loanword tensification is sensitive to the onset of the syllable following the tensification site (i.e. the assimilation effect). In addition, this trend in loanword tensification is likely to reflect the lexical distribution of Korean words. Then, to what acoustic property might be attributed the assimilation effect and the overrepresentation of two tense onset pairs in native word lexicon?

In order to explain this question, Ito (2014) hypothesized the long-distance creakiness assimilation between vowels in the relevant syllables. In Korean, a vowel followed by a tense consonant typically has a creaky voice. According to Ito's (2014) suggestion, the vowel in the word-initial syllable assimilates in creakiness to the one in the following syllable with a tense onset (e.g., CV. C'V̰ > CV̰.C'V̰). Following this, the lax consonant in the word-initial syllable is affected by the creaky voice of the following vowel, resulting in the tensification of the preceding consonant (e.g., CV̰.C'V̰ > C'V̰.C'V̰). In addition, Ito (2014) has suggested the diachronic sound change to support her explanation: e.g., /kos.kal/ > /ko.k'al/ > [k'o.k'al] 'Buddhist priest's hood', and /tuk.tə.pi/

---

are associated with the shortest VOT. Therefore, Oh (2009) concluded that some modulation of VOT by vowel context in English does not explain the rate of loanword tensification. However, the shortest VOT followed by high lax vowels are only observed in English voiceless stops (Nearey and Rochet 1994). Loanword tensification occurs when the target initial stops are voiced stops in English source words. On the other hand, for English voiced stops, Nearey and Rochet (1994) have observed that VOT duration of the initial stops becomes lonssible reason for the height effect of the loanword tensification.

> /tut.kə.pi/ (metathesis) > /tu.k'ə.pi/ > [t'u.k'ə.pi] 'toad' (Ito 2014: 389).

The present study supports Ito's (2014) account in that it can also be applied to the trends observed in loanword tensification. As previously mentioned, English prevocalic and word-final /s/ is consistently adapted as the Korean tense fricative /s'/ (Oh 1996, Y. Kang 2003, 2008). The onset of the vowel following the tense fricative /s'/ has either low or negative H1-H2 values, which indicates that it contains a creaky voice (Ito 2014, Kong et al. 2011). The creakiness of the vowel in the following syllable extends to the first syllable, and in turn, the creakiness of the preceding syllable triggers the tensification of the initial stop. For instance, the English word 'bus' is adapted as [pə.s'ɨ], where the vowel /ɨ/ in the second syllable contains the creaky voice. Following this, the creaky voice in /ɨ/ extends to the vowel in the initial syllable, resulting in [pə̰.s'ḭ]. Finally, the creakiness of /ə/ forces a preceding stop tensification on /p/, which is more harmonic with the creaky voice of /ə/ than /p/, resulting in [p'ə̰.s'ḭ] (e.g., /pə.s'ḭ/ > /pə̰.s'ḭ/ > [p'ə̰.s'ḭ] 'bus').

One might be concerned about the fact that there is a slight difference between the Yanbian and Seoul dialects of Korean, in that the former takes the creaky voice as a major distinctive cue to distinguish tense from lax consonants (Ito 2014), while in the case of the latter, the VOT parameter is sufficient to distinguish tense from lax stops (M. Kim 2004, Kong et al. 2011). Creaky voice as a major distinctive cue in Yanbian Korean is more likely to encourage the assimilation of the creaky voice from the following syllable's vowel to the vowel of the preceding syllable, which, in turn, triggers the regressive assimilation of two tense onset pairs. Nevertheless, the creakiness in the following vowel still plays a role in distinguishing tense from lax consonants in Seoul Korean (Oh and Yang 2013), therefore supporting the present study's assumption that the creaky voice may act as a trigger for the regressive assimilation of the two tense onset pairs in loanword tensification.

However, when we consider the articulatory movement of the assimilation effect of loanwords, it becomes quite tricky. Note that the tense consonants are typically characterized by the feature [+constricted glottis], and aspirated by [+spread glottis] (Halle and Stevens 1971, Hayes 2009). [+constricted glottis] is the opposite of [+spread glottis] in that the vocal cords make a narrow or closed glottis in the former, whereas in the latter, a wide glottis is produced (Hayes 2009). The extreme glottal status of both aspirated and tense consonants would make an articulatory effort greater than in lax consonants, which is laryngeally unmarked so that it needs less articulatory effort than tense consonants. For example, Ito (2014) has attributed the

co-occurrence restriction in aspirated-tense pairs to this difference in glottal states between aspirated and tense stops. In other words, it requires a significantly large amount of articulatory effort to pronounce two opposing glottal states, resulting in the co-occurrence restriction in aspirated-tense pairs.

Given this, the assimilation effect of loanword tensification also bears a great amount of articulatory effort, since the effect brings about two tense consonants within a word. That is, the effect requires much more articulatory effort than the original word having lax-tense pairs (e.g., more articulatory effort to pronounce [t'ɛns'ɨ] than [tɛns'ɨ] 'dance'). Thus, the assimilation effect seems unnatural respecting the articulatory movement. Even if there is no clear phonetic motive for the assimilation effect, it is one of the effects that shape the tendency in loanword tensification. This implies that the assimilation effect may be heavily dependent on the phonotactic knowledge of the Korean speaker.

5. Learning simulation

In this section, I will briefly demonstrate that the contextual distribution of the loanword tensification can be properly predicted by a machine learning simulation using the Optimality Theory (OT; Prince and Smolensky 1993/2004) constraints. Based on the observed contextual distribution of the loanword tensification, I set the following constraints.

(8) Constraint set adopted for the word-initial tensification of English loanwords
   a. IDENT-IO[-spread glottis, -constricted glottis] (henceforth, ID-IO[-s.g, -c.g])
      : The specification for the laryngeal feature [-spread glottis, -constricted glottis] of an input segment must be preserved in its output correspondent
   b. *TENSE/HIGH V
      : If the following vowel is high, tensification is blocked.
   c. *LAX/MONOSYLLABIC
      : If the word is monosyllabic, tensification occurs.
   d. C'-C'
      : If the second syllable onset has a [+constricted glottis] feature, word initial tensification occurs.

Following the standard OT, there should be only one optimal output which incurs the least serious violations of a set of constraints because these constraints should follow the concept of s*trict domination,* where the higher-ranked constraint in a pair of conflicting constraints takes precedence over the lower-ranked one. This results in a rigid ranking of constraints. Satisfaction of higher-ranked constraints has an uncompromisable priority over the satisfaction of lower-ranked ones. Thus, no possible degree of satisfaction of lower-ranked constraints can compensate for the violation of a single higher-ranked constraint. However, there are two (or more) optimal outputs in the variation data, which means that the phonological variation cannot be compatible with the concept of *strict dominance* or a strict ranking of the constraints. The tableau in (9) shows that the classic OT cannot account for the variation in the present study.

(9) Tableau of English loanword [pəs'ɨ] ~ [p'əs'i] 'bus' within the Classic OT frame (in case of the rigid ranking of ID-IO[-s.g, -c.g] >> C'-C')

Input: /bəs/	ID-IO[-s.g, -c.g]	C'-C'
☞ a. pəs'ɨ		*
☜ b. p'əs'ɨ	*!	

As we can see from the tableau above, the optimal output should be only [pəs'ɨ], which satisfies the highest ranking constraint, and there is no possibility that [p'əs'ɨ], which violates the highest constraint, can be the optimal output. This means that there should be no variation.

As discussed in Coetzee and Pater (2011), various modified theories and models have been suggested in order to explain the variation in the phonological pattern, including Partially Ordered Constraint (Kiparsky 1993, Anttila 1997), stochastic Optimality Theory (Hayes and MacEachern 1998, Hayes and Londe 2006), Noise Harmonic Grammar (Boersma and Pater 2008), Gradual Learning Algorithm (Boersma and Hayes 2001), and Maximum Entropy model (Goldwater and Johnson 2003). All these theories and models are based on the concept of assigning different weight values to the constraints in order to account for the non-categorical variation. The present study adopted the Maximum Entropy model (Goldwater and Johnson 2003), which is a statistical model. With this model, constraint weights are additive in log probability, resulting in an outcome where many violations of lower-ranked constraints may outweigh fewer violations of higher-ranked constraints. The software

reflecting the Maximum Entropy model, Maxent Grammar Tool (Hayes et al. 2009), was utilized in order to find the gradient constraint weighting. The judgment test data on English loanwords were learned with a default setting (μ=0; σ^2=100,000). The constraints and their calculated weights are shown in Table 26.

Table 26. Constraint weights learned from loanword data

Constraint	Weight
*TENSE/HIGH V	3.33
*LAX/MONOSYLLABIC	1.99
C'-C'	1.83
ID-IO[-s.g, -c.g]	1.76

The results show that the constraint with the highest weight is *TENSE/HIGH V. This explains the few cases of loanword tensification before high vowels. The constraint weight of *LAX/MONOSYLLABIC is higher than the weight of ID-IO[-s.g, -c.g] which captures monosyllabic word being more likely to undergo loanword tensification. In addition, higher constraint weight of C'-C' than that of ID-IO[-s.g, -c.g] suggests that the loanword tensification is more likely to occur when the onset of the syllable following the tensification site is a tense /s'/. Although the constraint weight of *LAX/MONOSYLLABIC and C'-C' is higher than that of ID-IO[-s.g, -c.g], *TENSE/HIGH V is much higher than any other constraint weights. Given this, we can predict that the rate of loanword tensification will be low when the height of the vowel following the tensification site is high, even if the word is monosyllabic or has tense fricative in the second syllable. Table 27 shows the predicted rates after learning with the constraint weights.

Table 27. Observed and learned loanword tensification rates; frequencies of each category are in parenthesis
(H: high, NH: non-high, S: sonorant, NT: no-tense, T: tense)

Input	Candidate	OBS. %	PRED. %	Example	Gloss
C.H.S	C.H.S	100 (1,240)	99.4	pi.ə	beer
	C'.H.S	0 (0)	0.6	-----	-----
C.H.NT	C.H.NT	98.1 (844)	99.4	pi.si.kʰit	biscuit
	C'.H.NT	1.9 (16)	0.6	p'il.tiŋ	building
C.H.T	C.H.T	100 (120)	96.3	cu.s'ɨ	juice
	C'.H.T	0 (0)	3.7	-----	-----

C.NH.S	C.NH.S	86.5 (1,177)	85.3	te.i	day
	C'.NH.S	13.5 (183)	14.7	k'e.im	game
C.NH.NT	C.NH.NT	84.6 (1,556)	85.3	te.sɨ.k^{h}ɨ	desk
	C'.NH.NT	15.4 (284)	14.7	c'e.cɨ	jazz
C.NH.T	C.NH.T	46 (92)	48.2	te.s'ɨ	death
	C'.NH.T	54 (108)	51.8	p'ak.s'ɨ	box
C.H#	C.H#	95 (190)	95.7	puk	book
	C'.H#	5 (10)	4.3	p'i	B
C.NH#	C.NH#	44.7 (170)	44.4	kən	gun
	C'.NH#	55.3 (210)	55.6	k'ol	goal

The Coefficient of Determination between the observed and the predicted rate was quite high ($R^2 = 0.99$), which means that the constraints applied, and their weights accurately predicted the variation trends in the English loanwords. However, it is difficult to apply the same constraint set to the native tensification data since the data turned out to have only the height effect among the three phonological factors. For this reason, it is not possible to compare the ratios observed in the English loanword data with the expected rate in the Korean word-initial tensification data. Thus, the present results merely show that the tensification of English loanwords is affected by the grammar with the constraint set and their weights.

6. Conclusion

In this study, I have explored the variable tensification of word-initial stops observed in English loanwords as they are adapted in Korean. This study is based on the idea that the loanword tensification reflects the distribution of native tensification or extant patterns of the native word lexicon. In order to verify this idea, I first investigated the contextual distribution of loanword tensification from two different sets of data; a judgment test and newspapers from the 1890s to 1950s. From the survey results of these two different data, I confirmed that there are three statistically significant phonological factors that shape tendencies in loanword tensification. The trends showed that loanword tensification is more likely to occur when the height of vowel following the tensification site is non-high (i.e. the height effect), when the word is monosyllabic (i.e. the length effect), and when the onset of the syllable following the tensification site has a phonation type of tense (i.e. the assimilation effect).

Given the contextual distribution of loanword tensification, I examined the two

possible sources, native tensification and native word lexicon, which are thought to affect the trends in loanword tensification. The survey results of the two possible sources showed that the effects of loanword tensification partially reflected the salient distribution of both sources. The height effect of the loanword tensification was also confirmed in native tensification. On the other hand, the height, length, and assimilation effect of loanword tensification mirror the extant distribution of Korean nouns.

In addition, it has been discussed that the phonetic property that motivates each contextual factor found in loanword tensification, focusing on the acoustic and articulatory property of each context. For the height effect of loanword tensification, the relation between the VOT value of the word-initial stop and the following vowel's height was considered. Previous studies have shown that the VOT values of word-initial stops become longer before high vowels, and in contrast, become shorter before non-high vowels, implying that the height effect is supported by a phonetic motive. Regarding the assimilation effect, it is greatly affected by the salient lexical distribution of Korean nouns, even though the effect requires an additional articulatory effort.

Finally, I performed a learning simulation for the trends observed. To calculate the weights of each constraint, the concept of Maximum Entropy model (Goldwater and Johnson 2003) was adopted. It was confirmed that the distribution observed in English loanwords is accurately predicted when it is learned with the proper constraints and their associated weights.

This study offers an observational contribution in that new factors affecting loanword tensification were found, premised on the idea that loanword phonology reflects native phonology. In addition, it firstly explores the question of what phonological aspects of native language drive the contextual distribution of loanword tensification, which has not been discussed in detail in previous studies.

REFERENCES

AHN, SANG-CHEOL. 2003. English interdental substitution. *English Language and Literature* 49.5, 981-1004.

ANTTILA, ARTO. 1997. Deriving variation from grammar: a study of Finnish genitives. In F. Hinskens, R. van Hout, and L. Wetzels (eds.). *Variation, Change, and Phonological Theory* 35–68. Amsterdam: John Benjamins.

BATES, DOUGLAS, MARTIN MAECHLER, and BIN DAI. 2008. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-28. http://lme4.r-forge.r-project.org/

BOERSMA, PAUL and BRUCE HAYES. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32, 45-86.

BOERSMA, PAUL and JOE PATER. 2008. Convergence properties of a gradual learner in harmonic grammar. *Rutgers Optimality Archive*, 970.

COETZEE, ANDRIES and JOE PATER. 2011. The place of variation in phonological theory. In J. Goldsmith, J. Riggle, and A. Yu (eds.). *The Handbook of Phonological Theory* (2nd Edition). Malden, MA and Oxford, UK: Blackwell. 401-434.

DAVIS, STUART and MI-HUI CHO. 2006. Phonetics versus phonology: English word final /s/ in Korean loanword phonology. *Lingua* 116.7, 1008-1023.

DOCHERTY, GERARD. J. 1992. *The Timing of Voicing in British English Obstruents.* (Vol. 9). Dordrecht, The Netherland: Walter de Gruyter.

GALLAGHER, GILLIAN. 2010. Perceptual distinctness and long-distance laryngeal restrictions. *Phonology* 27, 435-480.

GOLDWATER, SHARON, and MARK JOHNSON. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm Workshop on Variation Within Optimality Theory*, 111-120. Stockholm: Stockholm University.

HALLE, MORRIS and KENNETH N. STEVENS. 1971. A note on laryngeal features. *MIT Quarterly Progress Report* 101, 198-212.

HAN, MYEONG-SUK. 2010. *A Study on the Word-initial Glottalization in Korean*. PhD Dissertation. Kon-Kuk University.

HAYES, BRUCE. 2009. *Introductory Phonology*. Chichester: John Wiley & Sons.

HAYES, BRUCE, and ZSUZSA CZIRÁKY LONDE. 2006. Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology* 23, 59-104.

HAYES, BRUCE, and MARGARET MACEACHERN. 1998. Quatrain form in English folk verse. *Language* 74, 473-507.

HAYES, BRUCE, COLIN WILSON, and BENJAMIN GEORGE. 2009. Maxent grammar tool [Computer program]. http://www.linguistics.ucla.edu/people/hayes/Maxent-GrammarTool.

HEO, WOONG. 1965 *Kukeuemunhak* [Korean Phonology]. Cengumsa.

HERD, JONATHON. 2005. Loanword adaptation and the evaluation of similarity. *Toronto Working Papers in Linguistics* 24, 65-116.

ITO, CHIYUKI. 2014. Compound tensification and laryngeal co-occurrence restrictions in Yanbian Korean. *Phonology* 31.3, 349-398.

ITO, CHIYUKI, YOONJUNG KANG and MICHAEL KENSTOWICZ. 2006. The adaptation of Japanese loanwords into Korean. *MIT Working Papers in Linguistics* 52, 65-104.

KANG, BUM-MO and KIM, HEUNG-KYU. 2009. *Frequency of Korean*. Seoul: Hankookmunhwasa.

KANG, HIJO and OH, MIRA. 2015. Dynamic laryngeal co-occurrence restrictions in Korean. *The Phonology-Morphology Circle of Korea Winter Conference 2015*, 37-38.

KANG, YOONJUNG. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20, 219-273.

_____. 2008a. Tensification of voiced stops in English loanwords in Korean. *Harvard Studies in Korean Linguistics XII*, 179-192.

_____. 2008b. The adaptation of English /s/ in Korean. *Inquiries into Korean Linguistics* 3, 1-13.

_____. 2011. Loanword phonology. *The Blackwell Companion to Phonology* 4, 1003-1026.

KAWAHARA, SHIGETO, HAJIME ONO, and KIYOSHI SUDO. 2006. Consonant co-occurrence restrictions in Yamato Japanese. *Japanese/Korean Linguistics* 14, 27-38.

KIM, AYOUNG. 2009. *(The) Phonological Study of the Formation of Loanwords*. MA Thesis. Seoul National University.

KIM, KYU-SUN. 1976. A study on the influence of Japanese and Japanese loanwords on the formation of Korean loanwords. *The Korean Language and Literature* 35, 19-39. Hankoogeomunhakhoe [The Society of Korean Language and Literature].

KIM, MIDAM. 2004. *Correlation between VOT and F0 in the Perception of Korean Stops and Affricates*. MA Thesis. Seoul National University.

KIM, NA-YEON. 2012. *A Study on Word-initial Glottalization in English Loanwords*. MA Thesis. Pusan National University.

KIM, SOO-HYUN. 2003. An analysis of writing loanwords in The Dictionary of Modern-Joseon-loanwords. *Ewha Eomun Noncip* 21, 249-271. Ewha Womans University.

KIM, SUNG-KYU. 2001. Strategies of spoken and written language. *Eoneohag* 30, 65-88. The Linguistic Society of Korea.

KIM-RENAUD, YOUNG-KEY. 1974. *Korean Consonantal Phonology*. PhD Dissertation. University of Hawaii.

KIPARSKY, PAUL. 1993. An OT approach to phonological variation. *Handout from Rutgers Optimality Workshop 1993*.

KLATT, DENNIS. 1975. Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research* 18, 686-706.

KONG, EUN JUNG, MARY. E BECKMAN and JAN EDWARDS. 2011. Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics* 39.2, 196-211.

KUBOZONO, HARUO. 2006. Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua* 116.7, 1140-1170.

KWON, YOUNG-JU. 1995. *A Study on the Difference between the Standard Writing and the Customary Writing*. MA Thesis. Sookmyung Women's University.

LEBEN, WILLIAM R. 1973. *Suprasegmental Phonology*. PhD Dissertation. MIT.

LEE, HYUN-BOK. 2002. *Standard Korean Pronouncing Dictionary*. SNU Press.

LEE, MI-JAE. 1993. New interpretation on intensification. *Malsori* 25.1, 27-37. The Phonetic Society of Korea.

_____. 1989. Socio-linguistic research on the initial intensification in Korean. *Thesis Collection* 7, 61-73. The University of Suwon.

LEE, SEUNG-CHULL. 1982. A study on tensification. *UIT Report* 13.1, 227-231.

LEE, SHIN-SUNG. 1981. *A Study of Korean Consonants*. MA Thesis. Kon-Kuk University.

LUKE, KANG-KWONG, and CHAAK-MING LAU. 2008. On loanword truncation in Cantonese. *Journal of East Asian Linguistics* 17.4, 347-362.

NATIONAL INSTITUTE OF KOREAN LANGUAGE. 2007. *Oelaeeopyogibeop*. National Institute of Korean Language.

NEAREY, TERRANCE M., and BERNARD L. ROCHET. 1994. Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of International Phonetic Association* 24.1, 1-19.

OH, MIRA. 1996. An optimal analysis for Korean loanword phonology. *Journal of English Linguistics* 1, 157-174.

_____. 2004. English stop adaptation as output-to-output correspondence. *Phonological Studies* 7, 165-172. The Phonological Society of Japan.

_____. 2009. The phonetics and phonology of English voiced stop adaptation in

Korean, *Proceedings of the 2nd International Conference on East Asian Linguistics*, Simon Fraser University, Vancouver, British Columbia.

OH, MIRA and HUI YANG. 2013. The production of stops by Seoul and Yanbian Korean speakers. *Journal of the Korean Society of Speech Sciences* 5.4, 185-193. The Korean Society of Speech Sciences.

OHALA, J. JOHN. 1976. A model of speech aerodynamics. *Report of the Phonology Laboratory (Berkeley)* 1, 93-107.

_____. 1981. Articulatory constraints on the cognitive representation of speech. *Advances in Psychology* 7, 111-122.

PARK, DONG-GEUN. 2000. Word-initial tensification in Seoul Korean. *Eoneohak* 27.1, 179-200. The Linguistic Society of Korea.

PARK, JUNGJIN. 2008. A study on the fortis in the orthography of loanwords. *Grammar Education* 9, 185-205.

PRINCE, ALAN, and PAUL SMOLENSKY. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell. Originally circulated. 1993 as Technical Report TR-2. Rutgers Center for Cognitive Science/ Technical Report CU-CS-696-93. University of Colorado at Boulder Department of Computer Science.

R CORE TEAM. 2014. *A language and environment for statistical computing* (Version 3.1.2) [Computer program]. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

ROSE, YVAN and KATHERINE DEMUTH. 2006. Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua* 116, 1112–1139.

SHIN, SEUNG-HOON and STUART DAVIS. 2004. Where have all the lax stops gone?: on the possible restructuring of the Korean stop system. In *Japanese/Korean Linguistics* 13, Stanford CA: CSLI.

SILVA, DAVID J. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23.2, 287-308.

YEO, SANG-PIL. 1985. Natural processes and fortition. *Hyotay Nonmuncip* 31, 89-104. Catholic University of Taegu-Hyosung.

YU, SUNG-KON. 1987. English loanwords usage in the Korean language. *Dongseomunhwa* 19, 287-307. Keimyung University.

ZURAW, KIE. 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *NLLT* 28.2, 417-472.

Hyoju Kim
Department of Linguistics
Seoul National University
Daehak-dong, Gwanak-gu, Seoul
Korea 151-745
e-mail: kimhj14@snu.ac.kr

received: March 8, 2016
revised: August 4, 2016
accepted: August 13, 2016