Loss of productivity in phonological processes:

The case of Korean vowel harmony

Jinyoung Jo

Contents

1 Introduction

A phonological process can become less productive over time. When a phonological rule begins to lose productivity, applying less and less frequently, what might be the cause?[1] Previous studies suggested factors such as sound change (Dresher, 1985; S. N. Lee, 1947), opacity (Kiparsky, 1971; O'Bryan, 1974) and analogical leveling (Hock, 1991; King, 1969); still other factors will be discussed below. Moreover, we are also interested in details of the breakdown. Which element of the language is affected initially while others are more resistant to the breakdown, and why should such differences arise? Studying how and why rules retreat is an important issue in phonology, as it has implications for phonological learnability and language acquisition (Foulkes & Vihman, 2015; King, 1969; Lightfoot, 1999; Sen, 2016).

The present study concerns the case of Korean vowel harmony, a process that is arguably losing its productivity, and explores how and why the process is receding. Specifically, this study aims to contribute to understanding the process of rule breakdown by examining the harmony pattern of suffix alternation in different registers of speech and by investigating phonological and morphological factors that modulate harmony. A corpus study of the harmony process revealed a nuanced pattern: it has begun to weaken in a relatively casual style of speech, and the non-harmonized form is more frequently found in particular phonological and morphological contexts than in others.

It has been argued that a receding phonological rule often goes through the stage of being confined to derived environments (Kiparsky, 1973). That is, it is common for a phonological rule to apply initially both within morphemes and across morpheme boundaries, but later operate only across morpheme boundaries. Korean vowel harmony may be considered such a case. In Middle Korean, it applied both root-internally and across morpheme boundaries, while in Contemporary Korean, it still applies in verbal inflections but the trace of harmony in the lexicon is relatively weak. In recent research, the concept of the derived environment rule

[1] Throughout this paper I will use "rules" to designate morphophonemic processes, but this is not meant to be a theoretical commitment. The same analysis could in principle be applied to either a constraint-based or rule-based framework.

has itself become more nuanced, with a gradient connection between alternations and stem behavior. Chong (2019) found that when *quantitative* patterns of alternation and phonotactics are considered, a derived environment effect might actually have statistical support in the lexicon. He further claims that it is hard to learn alternation that is not supported by phonotactic generalizations and that alternation is productive to the extent that patterns in the lexicon support it. With this theoretical background, an important focus of the present study concerns the extent to which the lexicon and the alternation (mis-)match in terms of the harmony pattern, as it is imaginable that loss of productivity in Korean vowel harmony in suffix alternation may be attributed to lack of phonotactic support.

In this paper, I first present evidence that Korean vowel harmony is receding. I then test the hypothesis that such loss of productivity may be ascribed to an incongruity between alternation and phonotactics by quantitatively investigating the degree to which they mismatch. In line with Chong's (2019) claim, Korean vowel harmony turns out to be a case where an alternation with little phonotactic support is losing productivity.

1.1 Derived environment effects

Let us consider derived environment rules in more detail. A well-known example is Korean palatalization. In Korean, /t^h, t/ palatalize to [c^h, c] before [i] across a morpheme boundary (1a, c), while [t^{h}i] and [ti] sequences are attested within stems (1b, d).

(1) *Palatalization applies only in derived environments*

a. /pat^h-i/	[pac^h-i]	'field-NOMINATIVE'
b. /t^{h}i/	[t^{h}i]	'dust, flaw'
c. /mat-i/	[mac-i]	'eldest-NOMINALIZER'
d. /mati/	[mati]	'knot'

Previous studies of such derived environment effects, also known as non-derived environment blocking, have sought to characterize rules that apply only across morpheme boundaries or when fed by another phonological rule. Kiparsky (1973) argued that the rules which ap-

ply only in derived environments are non-automatic neutralizing rules. Mascaro (1976) appealed to the Strict Cyclic Condition, arguing that cyclic rules only apply in derived environments. Kiparsky (1993) claimed that derived environment effects arise as the result of structure-building rules applying to underspecified representations. Within the framework of Optimality Theory (Prince & Smolensky, 1993, 2004), Łubowicz (2002) adopted a constraint conjunction approach, while McCarthy (2003) introduced comparative markedness constraints.

Paster (2013) provided a diachronic account for a derived environment rule, arguing that alternation and phonotactics are free to differ with regard to a phonological generalization. The claim is that if alternation and phonotactics happened to be affected by the same factors diachronically, they will match; if they have undergone changes independently of each other, they will not.

In contrast, Chong (2019) suggested that there is in fact a link between phonotactics and alternation in several well-known cases of derived environment effects. For instance, Korean palatalization presented in (1a, c) is an alternation supported by a gradient phonotactic constraint by which [t^{h}i] and [ti] are severely underrepresented in the lexicon, although not totally illicit (see 1b, d). Further, a series of artificial grammar learning experiments suggested that alternation learning is facilitated by phonotactic learning and learners of an alternation-phonotactics mismatch language fail to learn the alternation. Based on these findings, Chong argued that alternation is productive to the extent that it is supported by phonotactic generalizations.

While Chong studied the link between alternation and phonotactics using rigorous statistical methods, the claims have not been tested in a sufficient number of cases. As he notes, it is necessary to consider quantitative patterns of alternation and phonotactics in other cases described as derived environment effects, as each case might differ in terms of the degree to which phonotactic generalizations match alternation patterns and the extent to which alternation is productive, affected by the strength of alternation-phonotactics link.

1.2 Prospectus

This study addresses two research issues. First, it provides documentation for the retreat of Korean vowel harmony. The evidence comes from a variety of sources including newspaper articles, a spoken corpus, and two wug tests reported in previous studies (H. Jang, 2017; H. Kang, 2012). The frequency of rule application in these sources shows that the harmony process is undergoing a typical pattern of rule loss, retreating first in the vernacular style of speech (Labov, 1966, 1972; Maclagan, 2000; Rohena-Madrazo, Simonet, & Paz, 2006). Second, the paper addresses the issue raised by Chong (2019), that is, how the degree to which phonotactics matches alternation affects the productivity of the alternation. An investigation of the lexicon shows that Korean vowel harmony is indeed a derived environment rule, evidenced by observation that phonotactics supports alternation in only an indirect way. These findings suggest that lack of phonotactic support for alternation might help explain the loss of productivity of Korean vowel harmony.

The remainder of the paper is organized as follows. Section 2 describes the pattern of Korean vowel harmony, focusing on its productivity and variability. Section 3 examines variable patterns of harmony in suffix alternation in two registers that differ in formality, followed by a comparison between the patterns of harmony in alternation and those in the lexicon (Section 4). Section 5 then reviews the results of nonce word tests conducted in previous studies (H. Jang, 2017; H. Kang, 2012). Finally, Section 6 summarizes the findings of this study and discusses how productivity of Korean vowel harmony might be affected by the relation between alternation patterns and relevant phonotactic restrictions.

2 Background: Korean vowel harmony

2.1 Basic patterns of Korean vowel harmony

I begin by describing basic patterns of Korean vowel harmony. In Korean verbal inflection, certain vowel-initial suffixes have two allomorphs, one of which has [a] as its first vowel and the other [ʌ]. The selection of the suffix vowel is determined by the final vowel of the stem. The initial suffix vowel is [a] if the last stem vowel is /a/ or /o/ (2a); otherwise, it is [ʌ] (2b). (The uppercase A represents the underlying form of the alternating suffix vowel.)

(2) *The basic pattern of vowel harmony*

		Underlying form	**Surface form**	**Last stem vowel**	**Gloss**
a.	i.	/pat-AsʌA/	[pat-asʌ]	a	'to receive-because'
	ii.	/cop-A/	[cop-a]	o	'to be narrow-DECLARATIVE'
b.	i.	/cʌp-A/	[cʌp-ʌ]	ʌ	'to fold-DECLARATIVE'
	ii.	/se-Ala/	[se-ʌla]	e	'to count-IMPERATIVE'
	iii.	/kʌnil-A/	/kʌnil-ʌ/	i	'to stroll-DECLARATIVE'
	iv.	/mantɨl-Ala/	[mantɨl-ʌla]	ɨ	'to make-IMPERATIVE'
	v.	/mul-AsʌA/	[mul-ʌsʌ]	u	'to bite-because'

The chart in (3) provides a vowel inventory of Korean, following H. Kang (2012) and Shin, Kiaer, and Cha (2013), among others. It assumes seven contrasting vowel phonemes. This represents, I believe, a vocalic system that holds good for the majority of current speakers of Korean. However, Korean dialects differ in their vowel systems, and more conservative dialects can have eight or even ten contrasting vowels (Y.-m. Y. Cho, 2016; Shin, 2015). These additional vowels (/ɛ/, /ø/, /y/) are spelled distinctly in the standard orthography which is used as the system of transcription in the Sejong corpus employed in this study. In analyzing the corpus, I have mapped all of these distinct spellings onto their normal outcomes in seven-vowel varieties of Korean. The intent is to give the most likely guess for what spoken vowels were intended by the speakers and writers whose output formed the basis of the corpus.

(3) *A vowel inventory of Korean (shaded cells = [ATR] vowels)*

	front	back	
	unrounded	unrounded	rounded
high	i	ɨ	u
mid	e	ʌ	o
low		a	

We can notice in the vowel chart that it is difficult to find a phonetically transparent distinctive feature which can separate the two harmonic groups of vowels, i.e. /i, ɨ, u, e, ʌ/ and /a, o/. The vowels /a, o/, traditionally termed 'light' vowels, do not form a natural class; nor do the 'dark' vowels /i, ɨ, u, e, ʌ/.[2] Some studies propose backness as the harmonic feature (C.-W. Kim, 1973) while others suggest lowness (Ahn, 1985; J. S. Kim, 2007; McCarthy, 1983). A greater number of studies adopt [± ATR] or [± RTR] (M.-H. Cho, 1994; Chung, 2000; Hong, 2008; H. Kang, 2012; Y.-S. Kim, 1984; B.-G. Lee, 1985; J.-S. Lee, 1992). For concreteness, I will follow the latter set of studies and adopt [ATR] (for /i, ɨ, u, e, ʌ/) and [RTR] (for /a, o/) as the harmonic feature in the current study, noting that this choice remains controversial on phonetic grounds.

As noted above, vowel harmony in Contemporary Korean is arguably a case of a derived environment rule. The harmony rule of suffix alternation has been studied extensively in the literature (Ahn, 1985; M.-H. Cho, 1994; S.-M. Cho, 2001; Hong, 2008; H. Kang, 2012; C.-W. Kim, 1973; J. S. Kim, 2007; Kim-Renaud, 1976, among others), and harmony applies fairly regularly in alternation. In stems, by contrast, disharmonic patterns are common, as illustrated by (4). An [ATR] vowel can be followed by a [RTR] vowel as in (4a-c), and a [RTR] vowel may be followed by an [ATR] vowel as in (4d-f).

[2] In the system of Korean ideophones /a, o/ are semantically associated with being light and bright, while /i, ɨ, u, e, ʌ/ are associated with being dark and heavy (Chae, 2003; K.-M. Lee & Ramsey, 2011; Sohn, 1999).

(4) *Disharmony is common within morphemes*

a. [ʌmma] 'mom'

b. [nuna] 'older sister'

c. [selo] 'vertical direction'

d. [namʌci] 'remainder'

e. [kaɨl] 'autumn'

f. [sonyʌ] 'girl'

Phonotactic generalizations in the lexicon that might reflect the harmony pattern of verbal inflection have not been studied extensively. A few studies suggest that the Korean lexicon has gradient phonotactic restrictions that reflect the harmony pattern of ideophones (Hong, 2010; Park, 2020). However, the grouping of harmonic vowels observed in ideophones is different from that of verbal inflections (see Hong (2010) and Jun (2018) for details). The present study aims to investigate the vowel co-occurrence restrictions in the lexicon that directly reflect the harmony pattern of verbal inflection, following the research program of Chong (2019) (Section 4). It also explores whether those phonotactic restrictions are conditioned by the same factors that modulate the variable patterns of harmony in alternation (variation in harmony will be discussed below in Section 2.3).

2.2 Productive harmony in Middle Korean

For purposes of documenting gradual retreat of vowel harmony, it is useful to review vowel harmony in Middle Korean, a stage at which the process was still fully productive. In Middle Korean, the ancestor of the present-day Korean spoken from the 10th century to the 16th century, the vowel system was symmetrical with regard to the number of vowels that belonged to each harmony group. The discussion of this section follows S. N. Lee (1947), W. Kim (1971), W. Kim (1978), K.-M. Lee (1998), Baek (1999), Sohn (1999) and H. Kang (2012). In Middle Korean, in which vowel harmony applied regularly both within (native) roots and across morpheme boundaries, the alternating vowel pairs were [a]~[ʌ], [ə]~[ɨ] and [o]~[u], and [i] was neutral. The

vowels [a], [ə] and [o] were the 'light' vowels, and [ʌ], [ɨ] and [u] were the 'dark' vowels. (Note that among these vowels, [ə] does not exist in Contemporary Korean, as will be explained below.) Some researchers argue that the harmonic feature was [back] with a corresponding vowel system in Figure 1 (a), while others argue that it was [RTR] and that the vowel system can be reconstructed as in Figure 1 (b).

Figure 1: Reconstructed vowel systems of Middle Korean

In Middle Korean, vowel harmony applied in broader morphological and phonological environments. First, whereas vowel harmony in Contemporary Korean applies only across a verb stem and a suffix, it was obeyed even within (native) roots in Middle Korean, as shown in (5). A root consisted of vowels from the same harmonic group, with possible existence of the neutral vowel [i]. Note that the Contemporary Korean forms of each root in (5) contain vowels from the different harmonic groups. Moreover, while [a]/[ʌ]-initial suffixes are the only ones eligible for harmony in Contemporary Korean, suffixes that began with other vowels also participated in vowel harmony in Middle Korean, as illustrated in (6). For example, [o]-initial suffixes alternated with [u]-initial ones, as shown in (6a). The ACCUSATIVE case marker in (6c), which has the invariant form [-ɨl] in Contemporary Korean, alternated between [-əl] and [-ɨl] in Middle Korean to harmonize with stem vowels. We can see in (6c-d) that the domain of harmony in Middle Korean was broader compared to Contemporary Korean in that the harmony applied not only in suffixes that combined with verb stems (as shown in (6a-b)) but also in case markers that were attached to nouns. This is because case markers that were eligible for harmony in Middle Korean no longer meet the conditions for harmony in Contemporary Korean after undergoing a series of sound changes (which will be discussed shortly). Both in Middle Korean and in Contemporary Korean, if at least one of the allomorphs of a suffix or a case marker be-

gins with a consonant, it was not eligible for harmony. For instance, the VOCATIVE case marker alternated between [a] (after consonant-final noun) and [ja] (after vowel-final noun) with the invariant vowel.

(5) *Vowel harmony applied within morphemes in Middle Korean*

	Middle Korean	**Contemporary Korean**	**Gloss**
a.	[namo]	[namu]	'tree'
b.	[nilkup]	[ilkop]	'seven'
c.	[talə-]	[tali-]	'to be different'
d.	[toskapi]	[tok*epi]	'goblin'
e.	[kəəl]	[kaɨl]	'autumn'

(6) *More vowels participated in harmony in Middle Korean*

	Middle Korean	**Contemporary Korean**	**Gloss**
a.	[mak-o-]	N/A	'to block-will-'
	[mʌk-u-]	N/A	'to eat-will-'
b.	[mak-əni]	[mak-ɨni]	'to block-as'
	[mʌk-ɨni]	[mʌk-ɨni]	'to eat-as'
c.	[son-əl]	[son-ɨl]	'hand-ACC'
	[skum-ɨl]	[k*um-ɨl]	'dream-ACC'
d.	[saləm-əj]	[salam-ɨj]	'person-GEN'
	[kʌpup-ɨj]	[kʌpuk-ɨj]	'turtle-GEN'

However, a number of diachronic changes led to the disruption of the vowel harmony system, including vowel loss, changes in vowel quality and the adoption of Sino-Korean words, among which I mention only a few below. Breakdown of the harmony system is most attributed to loss of [ə], which was replaced with [ɨ] in non-initial syllables and with [a] in initial syllables. Let us take [kaɨl] 'autumn' in (5e) as an example. The 16th-century form of this word was [kəəl]. In the 17th century, the second [ə] was replaced with [ɨ] to give rise to [kəɨl]. In the 18th cen-

tury, [ə] in the initial syllable was transformed into [a], hence the present form [kaɨl]. Crucially, the change from [ə] to [ɨ] disrupted the harmony, since the two vowels belonged to different harmony groups. Further, it seems that the loss of [ə] is most responsible for the weakening of vowel harmony in case markers, many of which showed [ə]~[ɨ] alternation. This is illustrated in (6c-d), which show that case markers that had two alternating forms (i.e. [ə]-initial and [ɨ]-initial ones) in Middle Korean invariably surface as [ɨ]-initial form in Contemporary Korean.[3]

It is assumed that the change in the vowel quality of /ʌ/ has initiated sound changes that ultimately led to the loss of /ə/ outlined above. By late Middle Korean, /ʌ/ had moved backwards in the vowel space to become a central vowel. As /ʌ/ moved back towards /ɨ/, /ɨ/ was shifted to the position of /u/, which in turn was shifted towards /o/. Then /o/ moved towards /ə/, whose phonetic status became unstable. Ultimately, /ə/ was lost and replaced with other vowels.

In addition, a significant amount of Sino-Korean vocabulary, which did not conform to the harmony, was brought into Korean. This is illustrated by the Sino-Korean word [sonyʌ] 'girl' in (4f), which contains two disharmonic vowels. As a result of the adoption of Sino-Korean words, Korean became abundant with forms that did not follow the harmony.

For these diachronic reasons, then, vowel harmony that was once productive in Middle Korean now applies in limited contexts in Contemporary Korean. Harmony is not fully active within morphemes, case markers do not harmonize to the vowels of noun stems, and only a restricted set of verbal suffixes undergo harmony. As already mentioned, this is because it happens to be the case that a few verbal suffixes are the only ones that meet the conditions for harmony, i.e. must be vowel-initial and begin with [a]/[ʌ].

In sum, vowel harmony was active both within morphemes and across morpheme boundaries in Middle Korean. From the viewpoint of Chong (2019), we could also say that alternation was productive because it had phonotactic support in the lexicon. After a number of sound changes took place, the phonotactic harmony was disrupted while still leaving some of the al-

[3]Although researchers do not discuss loss of [o]~[u] alternation as much as that of [ə]~[ɨ], it seems that [o]~[u] alternation is assumed to be affected by loss of [ə]~[ɨ] alternation. At some point in Middle Korean, [o] and [u] tended to be pronounced as unrounded, such that [o]~[u] became similar to [ə]~[ɨ], whose alternation was already disrupted. This led to loss of harmony for [o] and [u] (Y. Jang, 1993).

ternation behind. The following section describes vowel harmony in Contemporary Korean, focusing on its limited productivity.

2.3 Limited productivity of harmony in Contemporary Korean

As noted above, in Contemporary Korean, alternation exhibits harmonic patterns, while disharmony is observed within morphemes. However, there is evidence that even the harmony in alternation is not fully productive. This section focuses on the limited productivity of vowel harmony in verbal inflection in Contemporary Korean.

Several characteristics of Korean vowel harmony of the present day converge to demonstrate its limited productivity. First, as mentioned above, it applies in a small subset of verbal suffixes, i.e. [a] or [ʌ]-initial ones. Strikingly, the process is not iterative. For instance, the harmony does not apply to the second vowel of the suffix in [pat-asʌ] 'to receive-because' (2a.i.) and [mantɨl-ʌla] 'to make-IMPERATIVE' (2b.iv.); if it had applied iteratively, we would have expected *[pat-asa] and *[mantɨl-ʌlʌ], respectively. The latter of these examples also shows that only the last stem vowel can be the trigger and other vowels in the stem do not affect the harmony.

Moreover, the harmony pattern is not fully reproduced in wug tests reported in previous studies (which will be discussed in greater detail in Section 5). H. Kang (2012) conducted a production experiment using real and nonce stems to test the speakers' intuition on the harmony pattern. When the stem vowel was /a/ or /o/, the proportion of harmonic suffix forms was only 32.0% for nonce stems, as compared to 76.7% for real stems (with the /-A/ 'DECLARATIVE, INTERROGATIVE, IMPERATIVE (familiar)' suffix, which particularly lowers the rate of harmony application; see Section 3.4). Given that the harmony did not apply categorically even in real words (the variable pattern will be discussed shortly), it may not be surprising to see disharmonic forms in wug tests. It is striking that the dominant pattern in speakers' response for wug words was the disharmonic one.

Furthermore, the harmony is not always observed even in ordinary speech and variation is shown in the selection of the suffix vowel. Specifically, the [ʌ] suffix allomorph is often observed

where [a]-allomorph is expected. The following data in (7) are taken from the corpus survey reported below.

(7) *The variable pattern of vowel harmony*

a. /cap-A/	[cap-a]~[cap-ʌ]	'to catch-DECLARATIVE'
b. /mac-A/	[mac-a]~[mac-ʌ]	'to be correct-DECLARATIVE'
c. /pat-A/	[pat-a]~[pat-ʌ]	'to receive-DECLARATIVE'

Considering that [ʌ] shows a wider distribution than [a], and that suffixes surface with [ʌ] when not eligible for harmony (e.g. the second suffix in /pat-As*-A/ 'to receive-PAST-DECLARATIVE', which is not subject to harmony, surfaces as [ʌ], as in [pat-as*-ʌ], cf. *[pat-as*-a]), it is plausible to think of variation shown in (7) as a regularization process. Such variable patterning constitutes another piece of evidence showing that vowel harmony is gradually losing its productivity. In judgment surveys in which younger and older speakers' well-formedness ratings were compared (H. Kang, 2012, 2016), it was found that younger speakers were more flexible in allowing variation compared to older speakers, also suggesting a loss of productivity over time.

It is widely accepted that innovative forms arise first in more casual and vernacular styles of speech (Labov, 1966, 1972; Maclagan, 2000; Rohena-Madrazo et al., 2006). My own intuitive sense as a native Korean speaker is that disharmonic forms are attested in informal varieties of Korean, but rarely in formal registers. As far as I know, no previous work has systematically investigated the harmony pattern in different styles of speech. In this study, I compare data drawn from newspaper articles and speech to examine how much variation exists in each register and which factors condition the harmony. The corpus study of the suffix alternation shows that the harmony rule has begun to break down in speech. Looking ahead, by examining gradient vowel co-occurrence restrictions in the lexicon in Section 4, I suggest that such loss of productivity may be related to the fact that stem phonotactics matches the alternation pattern only weakly.

3 A corpus study of suffix alternation

To examine the pattern of alternation in the written and spoken registers of Contemporary Seoul Korean, I used the Sejong text corpus, downloaded from the database of the National Institute of the Korean Language.[4] The corpus employs Korean orthography, which is nearly phonemic, so it is possible to recover from the orthography how the stem vowel and the suffix vowel are realized. A subset of the corpus is morphologically tagged, among which newspaper articles (1,625,713 words) and speech data (805,652 words) were subject to analysis. The newspaper data are drawn from articles such as news reports and editorials that were published from 1990 to 2003. The spoken data, recorded between 2001 and 2005, featured many different types of speech styles, including lectures, TV shows, sermons, daily conversation and presentations. The majority of the speakers were in their 20s at the time of recording.

To construct the alternation data, I extracted words which consist of a verb or an adjective stem followed by a V-initial suffix that participates in harmony, using the tags present in the corpus. Verb and adjective stems in Korean behave similarly in many phonological processes including vowel harmony and both take the suffixes that are considered in this study. I consulted B.-M. Kang and Kim (2004) and H. Kang (2012) to identify the harmonizing suffixes, the list of which is provided in (8). The uppercase letter A was used to represent the harmonizing vowel that surfaces as either [a] or [ʌ]. Note that all the suffixes in the list begin with the harmony-target vowel; as mentioned earlier, suffixes that begin with a consonant do not undergo harmony.

It should be noted that one of the suffixes is less likely to be used in newspaper articles. Specifically, the /-A/ 'DECLARATIVE, INTERROGATIVE, IMPERATIVE (familiar)' (henceforth /-A/ 'DECL'[5]), which is reported to surface frequently as the disharmonic [-ʌ] when the stem vowel is /a/ or /o/ (H. Kang, 2012), is exclusively used in casual conversations. In the newspaper

[4] https://ithub.korean.go.kr/

[5] Note that there are two types of 'DECLARATIVE, INTERROGATIVE, IMPERATIVE' suffixes, i.e. /-A/ (familiar) and /-Ajo/ (honorific). Throughout the paper, when referring to /-A/, I gloss it as 'DECL' without explicitly specifying 'familiar', for the sake of simplicity (unless the familiar/honorific distinction is crucial). /-Ajo/, which is mentioned only a few times in this paper, is glossed as 'DECL (honorific)'.

corpus, only 8 out of 995 words were suffixed with the /-A/ 'DECL'. Thus, one of the important factors that discourage harmony in the spoken data could be the imbalanced distribution of this specific suffix. I investigate this possibility in Section 3.5.

(8) *Harmonizing suffixes (*[a]- *or* [ʌ]*-initial)*

Suffix	**Meaning**
-A	DECLARATIVE, INTERROGATIVE, IMPERATIVE (familiar)
-Ajo	DECLARATIVE, INTERROGATIVE, IMPERATIVE (honorific)
-A	CONNECTIVE
-Ala	IMPERATIVE
-As*	PAST
-As*ʌs*	used to
-Asʌ	because
-Asʌjo	because (honorific)
-Ata	and
-Ataka	and
-Ato	though, but
-Aja	should
-Ajaman	only if, should
-Ajaci	only if / will, should
-Ajacijo	will, should (honorific)

3.1 Newspaper corpus

I first examined all sequences of a stem vowel and a suffix vowel in the newspaper data. The table in (9) shows how the suffix vowel is realized depending on the stem vowel, based on 62,457 word tokens, 3,827 word types and 1,706 stem types found in the newspaper articles. Two kinds of harmony occurrence rates are provided. One is rate of harmony based on word tokens, which was calculated by dividing the number of tokens in which the stem took a harmonic suffix by the

total number of tokens employing either of the suffix allomorphs. The other is rate of harmony based on word types, which is the mean harmony rate of words that belong to each stem vowel context. The shaded cells indicate harmonic forms.

(9) *Realization of suffix vowels in the newspaper corpus (shaded cells = harmonic forms)*

			Suffix vowel				
			Token			Type	
			a	ʌ	% harmony	# of word types	% harmony
stem vowel	a (regular)		11052	1	99.99	632	99.9
	a (p-irregular)		4	84	4.5	19	4.5
	o (regular)		3841	0	100	287	100
	o (p-irregular)		68	29	70.1	17	32.4
	i		0	23767	100	1536	100
	e		2	4221	99.95	219	99.8
	ʌ		0	5628	100	331	100
	u		0	4764	100	379	100
	ɨ	/...aC$_0$ɨ/ → [...aC$_0$]	2599	4	99.8	71	95.8
		/...oC$_0$ɨ/ → [...oC$_0$]	1193	6	99.5	37	98.6
		Other /ɨ/-stems	0	5194	100	299	100

For every stem vowel, the harmony rate is either 100% or virtually so, except for p-irregular /a/-stems and /o/-stems; p-irregular stems are a class of verb and adjective stems whose stem-final /p/ surfaces as [w] when followed by a vowel-initial suffix, as illustrated in (10). The second column of forms, with the consonant-initial suffix /-tamjʌn/ 'if', are included to show that these are stems with underlying /p/, not /w/. As (10d-f) shows, p-irregular /a/- and /o/-stems typically take the disharmonic [ʌ] suffix forms, although the harmonic variant does arise (Hong, 2008; H. Kang, 2012).[6]

[6]As a native speaker of Korean, I have an impression that the harmonic [a]-forms for p-irregular stems are used by the older generation and somewhat unnatural to younger native speakers.

(10) *P-irregular* /a/*-stems and* /o/*-stems frequently take disharmonic suffixes*

		V-initial suffix (/-A/ 'DECL')	**C-initial suffix (/-tamjʌn/ 'if')**	**stem meaning**
a.	/tʌp-A/	[tʌ**w**-ʌ]	[tʌ**p**-tamjʌn]	'to be hot'
b.	/c^{h}up-A/	[c^{h}u**w**-ʌ]	[cu**p**-tamjʌn]	'to be cold'
c.	/pulʌp-A/	[pulʌ**w**-ʌ]	[pulʌ**p**-tamjʌn]	'to envy'
d.	/pankap-A/	[panka**w**-ʌ] (~[panka**w**-a])	[panka**p**-tamjʌn]	'to be nice to see'
e.	/welop-A/	[welo**w**-ʌ] (~[welo**w**-a])	[welo**p**-tamjʌn]	'to be lonely'
f.	/kwelop-A/	[kwelo**w**-ʌ] (~[kwelo**w**-a])	[kwelo**p**-tamjʌn]	'to be in distress'

The pattern is confirmed in the current data, such that the harmony rate is lower for p-irregular /a/-stems (4.5%) and p-irregular /o/-stems (32.4%), as compared to regular /a/-stems (99.9%) and regular /o/-stems (100%).

Another complication involves /ɨ/-final stems. In Korean verbal inflection, stem-final /ɨ/ is deleted when stems combine with V-initial suffixes, and the vowel that becomes the last stem vowel as a result of this /ɨ/-deletion process triggers the harmony. To illustrate, /camkɨ-A/ 'to lock-DECL' is expected to surface as [camk-a]. Previous research has noted that when /a/ or /o/ becomes the last stem vowel as a result of /ɨ/-deletion, the suffix vowel can be either harmonic or disharmonic, e.g. both [camk-a] and [camk-ʌ] occur (Jun, 2018), thus behaving similarly to the stems in which /a/ or /o/ is the last vowel underlyingly. This is confirmed in the current data, as a few disharmonic forms are attested for /...aC$_0$ɨ/ stems (N=4) and /...oC$_0$ɨ/ stems (N=6).

Finally, two instances of disharmonic forms for /e/-stems were found, both of which were [p*es-a] 'to take away-DECL'. The stem /p*es/ is a shortened form of /p*eas/, whose last stem vowel is /a/, which leads to speculation that the speakers' use of [a]-initial suffix is a tendency to harmonize the suffix vowel to the original form of the stem. As will be shown below, the spoken data also exhibited two tokens of /p*es/ combined with an [a]-initial suffix.

We have seen that the harmony rate is nearly 100% in newspaper articles, except for p-irregular /a/- and /o/-stems for which disharmonic forms were frequently found. We can thus

say that for the newspaper corpus, which exhibits one of the most conservative registers, harmony applies fairly regularly. The near-categorical harmony in the newspaper corpus stands in contrast to the harmony pattern shown in the spoken data, which involve more disharmonic forms even outside the exceptional cases of p-irregular stems, as will be outlined in the next section.

3.2 Spoken corpus

The harmony pattern of the spoken data of the Sejong corpus is provided in the table in (11), which is formatted in the same way as (9) above. The data is based on 21,796 word tokens, 1,803 word types and 775 stem types.

(11) *Realization of suffix vowels in the spoken corpus (shaded cells = harmonic forms)*

			Suffix vowel				
			Token			Type	
			a	ʌ	% harmony	# of word types	% harmony
stem vowel	a (regular)		4600	909	83.5	414	91.7
	a (p-irregular)		2	100	2.0	28	0.68
	o (regular)		2534	0	100	142	100
	o (p-irregular)		123	25	83.1	20	42.1
	i		0	5312	100	420	100
	e		3	712	99.6	63	96.8
	ʌ		0	3378	100	300	100
	u		0	1435	100	210	100
	ɨ	/...aC$_0$ɨ/ → [...aC$_0$]	53	3	94.6	18	86.1
		/...oC$_0$ɨ/ → [...oC$_0$]	50	0	100	5	100
		Other /ɨ/-stems	0	2557	100	183	100

A number of main trends seen in the data are as follows. First, as is widely noted in the litera-

ture (M.-H. Cho, 1994; Han, 2009; Hong, 2008; H. Kang, 2012; Y. Kang & Ryu, 2015; Kim-Renaud, 1976), disharmonic forms are most frequently observed for /a/-stems. Not only p-irregular /a/-stems have a near-zero rate of harmony (0.68%), but even regular /a/-stems combine with [ATR] suffix allomorphs in a fair number of forms. /o/-stems differed from /a/-stems in that regular stems showed a categorical harmony; all the disharmony cases involved p-irregular stems, suggesting that /o/-stems exhibit a categorical harmony if we set aside the special case of p-irregular stems. Therefore, the effect of stem vowel quality on harmony rate is evident in the spoken corpus, i.e. of the two [RTR] vowels, /o/ is a stronger trigger than /a/.

Recall from Section 3.1 that /a/-stems in the newspaper data that took disharmonic suffix forms were mostly p-irregular stems (this was true for /o/-stems as well). This contrasts with the behavior of /a/-stems in the spoken corpus, because even regular /a/-stems in the spoken data frequently combined with disharmonic suffix forms. This constitutes evidence showing that the harmony is more strictly observed in the newspaper corpus compared to the spoken corpus, as we see that the innovative disharmonic forms are attested in a broader domain in the spoken data.

It is noticeable that the variation occurs in only one direction; we find that [ʌ]-initial suffix is observed where [a]-initial suffix is expected, but the opposite case is rarely attested. The [ATR] stem vowels /i/, /e/, /ʌ/ and /u/ show a near-categorical harmony, as they almost always combine with [ʌ]-initial forms. Three exceptions are /e/-stem words, i.e. [pet^{h}-**a**s*-ʌjo] 'to spit-PAST-DECL (honorific)', [p*es-**a**] 'to take away-DECL (familiar)' and [p*es-**a**s*-ʌjo] 'to take away-PAST-DECL (honorific)'. As mentioned in the previous section, it is likely that the selection of [a]-initial suffix for /p*es/ was influenced by its original form /p*eas/; otherwise they seem to be mere speech errors. I do not discuss these exceptions any further and assume in the analysis that these stem vowels categorically conform to the harmony pattern.

To sum up, far more disharmonic suffix forms were found in the spoken data compared to newspaper articles of Korean. In the following analysis, I focus only on /a/- and /o/-stems in the spoken corpus, which are subject to exceptions and variation, and demonstrate several

phonological and morphological factors that condition the harmony pattern.

3.3 The behavior of individual stems

Which stems are responsible for the variation shown in /a/-stems and /o/-stems? This section explores the variable pattern of individual /a/-stems and /o/-stems, based on the harmony rate averaged across individual words (i.e. the rightmost column in tables in (9) and (11)).

Figure 2 shows the harmony rate of individual words with /a/ or /o/ stem vowels. Among the 604 words found in the spoken data, I excluded words with token frequency of 1 or 2 (N=344). It can be seen that the majority of the words showed a very high harmony rate (90-100%). We also see a small number of words to which harmony rarely applies (0-10%), and even smaller number of words that belong to each bin with intermediate harmony rate. Of the 260 words, 201 words had 100% harmony rate, 44 words an intermediate harmony rate and 15 words 0% harmony rate; therefore, Korean vowel harmony seems to show a mix of lexical variation (different stems behave differently) and free variation (the very same stem behaves in different ways).

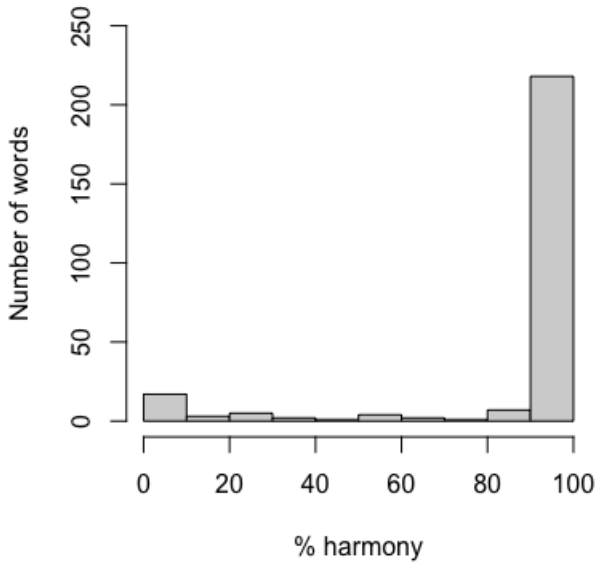

Figure 2: Harmony rates of individual words whose last stem vowel is [a] or [o]

Some example words that show 0%, intermediate, and 100% harmony rate are provided in (12)-(14). As mentioned in Section 3.2, it was found that p-irregular stems showed a very low

harmony rate, exemplified in (12a, b) and (13b).

(12) *Words with 0% harmony rate*

a. /kak*ap-A/	[kak*aw-ʌ]	'to be close-DECL'	(0%, N=7)
b. /komap-A/	[komaw-ʌ]	'to be thankful-DECL'	(0%, N=16)

(13) *Words with an intermediate harmony rate*

a. /mac-A/	[mac-a]~[mac-ʌ]	'to be correct-DECL'	(9.8%, N=569)
b. /welop-A/	[welow-a]~[welow-ʌ]	'to be lonely-DECL'	(14.3%, N=7)
c. /pat-A/	[pat-a]~[pat-ʌ]	'to receive-DECL'	(48.6%, N=35)
d. /anc-A/	[anc-a]~[anc-ʌ]	'to sit-CONNECTIVE'	(91.0%, N=111)

(14) *Words with 100% harmony rate*

a. /sal-Aja/	[sal-aja]	'to live-should'	(100%, N=25)
b. /salm-A/	[salm-a]	'to boil-CONNECTIVE'	(100%, N=7)
c. /nop^{h}-A/	[nop^{h}-a]	'to be high-DECL'	(100%, N=11)

To sum up, we found that the majority of words undergo harmony regularly, a few words never undergo harmony, and a fair number of words vacillate. In the next section, I investigate which phonological and morphological factors discourage harmony in the latter two cases.

3.4 The factors influencing variation

The variable harmony rate of /a/-stems and /o/-stems is conditioned by various phonological and morphological factors. This section focuses on five factors, i.e. the effect of suffix type, stem vowel quality, the number of intervening consonants between the two vowels, stem type (i.e. the effect of p-irregular stems already outlined above) and frequency of the stem. A subset of these factors were already investigated in Hong (2008) using the same data as this study, i.e. spoken data of the Sejong text corpus. He reported that p-irregular stems generally prefer disharmonic suffix forms and that regular stems show variation when (i) the last stem vowel is /a/ and (ii)

the stem ends in a consonant. I additionally show below that those two effects that emerged in regular stems are strongest for the /-A/ 'DECL' suffix.

3.4.1 The /-A/ 'DECL' suffix discourages harmony

We have seen that the newspaper corpus exhibited a higher harmony rate than the spoken corpus. In this section, I show that it was primarily the /-A/ 'DECL' suffix that was most responsible for the low harmony rate in the spoken data by comparing harmony rates in the three subcorpora, i.e. the newspaper corpus, the spoken corpus with the /-A/ 'DECL' suffix ("Declarative" dataset) and the spoken corpus with other suffixes ("Non-declarative" dataset). Recall that very few words in the newspaper corpus were suffixed with the /-A/ 'DECL' (see Section 3); therefore, any effect of this suffix would be minimal in the newspaper data.

It has been noted in H. Kang (2012) that the relative frequency of the disharmonic suffix form is the highest for the /-A/ 'DECL' suffix, the sentence-ending suffix that consists of only one vowel. The representative data are presented in (15); while the /-A/ 'DECL' frequently surfaces with a disharmonic vowel (15a-c), other suffixes show little variation, representative examples of which are shown in (15d-f). Note again in (15d-e) that harmony does not spread beyond the initial suffix vowel.

(15) /-A/ 'DECL' *is more likely to surface as the disharmonic form than other suffixes*

a. /c^{h}ac-A/	[c^{h}ac-a]~[c^{h}ac-ʌ]	'to find-DECL'
b. /al-A/	[al-a]~[al-ʌ]	'to know-DECL'
c. /nam-A/	[nam-a]~[nam-ʌ]	'to remain-DECL'
d. /c^{h}ac-As*-A/	[c^{h}ac-as*-ʌ]	'to find -PAST-DECL'
e. /al-Asʌ/	[al-asʌ]	'to know-because'
f. /nam-A/	[nam-a]	'to remain-CONNECTIVE'

The effect of suffix type on the harmony pattern in the current data is shown in Figure 3. In this figure and the following ones, the x-axis is the conditioning factor, and the y-axis represents the mean harmony rate of words that fall in each category of the independent variable. The

numbers at the top of each bar indicate the number of words that belong to that category. Confirming the finding of H. Kang (2012), the frequency of harmonic forms is substantially lower for the /-A/ 'DECL' suffix. The harmony rate for the Declarative portion of the spoken corpus was 43.8%, compared to 96.8% in the newspaper corpus and 93.2% in the Non-declarative spoken data. It should be noted that even the Non-declarative data had a lower harmony rate than the newspaper data, although the difference was quite small. As we will see in Section 3.5, this difference is in fact statistically significant. This suggests that there is at least a small effect of register, i.e. the harmony rate being higher in newspaper articles than in speech, that is independent of the effect of suffix type. In other words, the lower rate of harmony in the spoken data is not completely due to the effect of the /-A/ 'DECL'. [7]

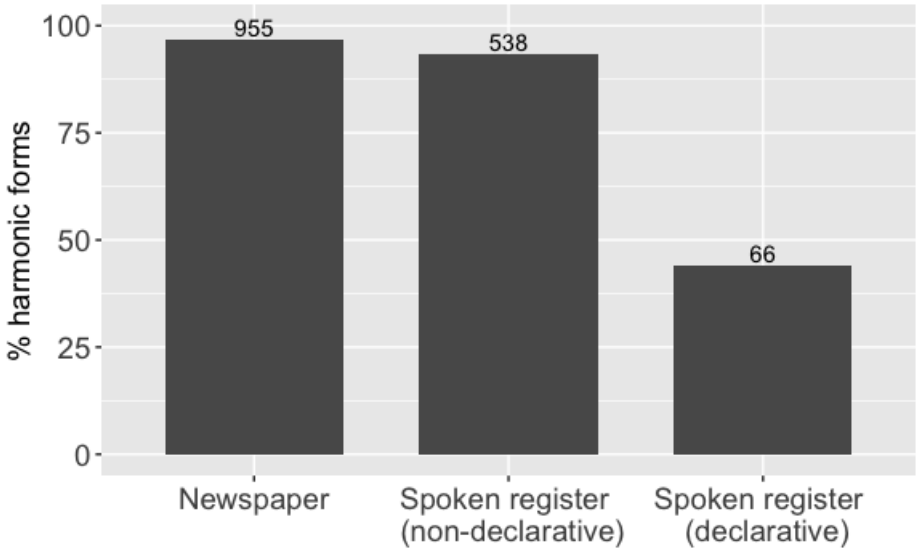

Figure 3: The effect of suffix type

[7] Among the twelve suffixes that belong to the Non-declarative category, some showed a categorical harmony, while others showed a very high yet not quite 100% harmony rate. In a logistic regression analysis implemented to test significance of each factor's effect on harmony rate in the spoken corpus (not reported here), the average harmony rate of the suffixes with a high but not 100% harmony rate was not significantly different from that of suffixes with 100% harmony.

3.4.2 /o/ is a stronger trigger than /a/

The quality of the last stem vowel conditions the harmonic pattern (M.-H. Cho, 1994; Hong, 2008; H. Kang, 2012; Kim-Renaud, 1976), illustrated in (16). If the final vowel of the stem is /a/, and the stem ends in a consonant, as in (16a-c), both the harmonic [a]-initial forms and the disharmonic [ʌ]-initial forms can surface. As we observed in (9) and (11), the stem vowel /o/ mostly takes [a]-initial suffix forms, illustrated in (16d-f).

(16) /a/*-stems combine with disharmonic suffixes while* /o/*-stems rarely do*

a. /pat-A/	[pat-a]~[pat-ʌ]	'to receive-DECL'
b. /mac-A/	[mac-a]~[mac-ʌ]	'to be correct-DECL'
c. /cap-A/	[cap-a]~[cap-ʌ]	'to catch-DECL'
d. /c*oc^{h}-A/	[c*oc^{h}-a]	'to follow-DECL'
e. /nol-A/	[nol-a]	'to play-DECL'
f. /nop^{h}-A/	[nop^{h}-a]	'to be high-DECL'

Figure 4 shows that the trigger vowel effect, in which /o/-stems more frequently take the harmonic [a]-initial suffix forms than /a/-stems, emerges only in the /-A/ 'DECL' suffix. In the newspaper data and the Non-declarative data, we hardly see a difference between /a/-stems and /o/-stems. It is striking to note that words whose stem vowel is /a/ and suffix is the /-A/ 'DECL' show a very low rate of harmony (28.9%)

As already mentioned in Section 3.2, it should be noted that all of the /o/-stems that take disharmonic forms were p-irregular stems. Therefore, if we only consider regular stems, /o/-stems should exhibit a categorical harmony. The effect of p-irregular stems will be presented in more detail in Section 3.4.4 below.

3.4.3 Hiatus

It has been reported that the variation arises only when the /a/-stems end in a consonant (M.-H. Cho, 1994; H. Kang, 2012). In other words, if there is no intervening consonant between the

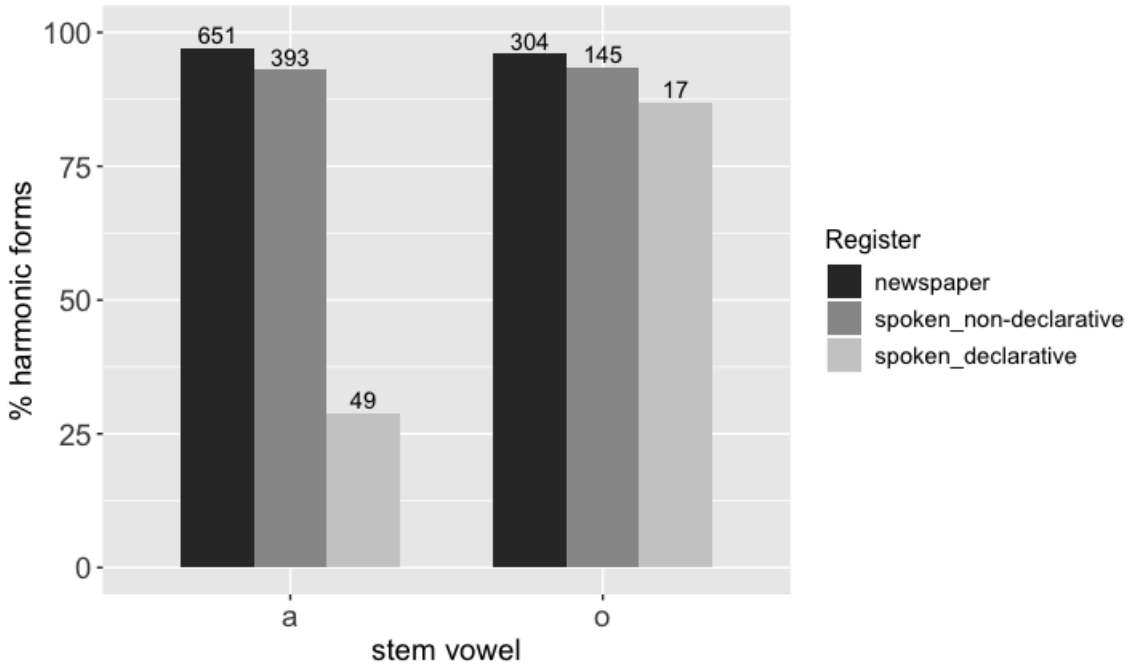

Figure 4: The effect of stem vowel

last stem vowel /a/ and the initial vowel of the suffix in the surface form, only the harmonic [a]-initial form is possible.

(17) *Previous research reports no variation in hiatus context*

a. /po-A/ [po-a] 'to see-DECL'

b. /nah-A/ [na-a] 'to produce-DECL'

c. /s*ah-A/ [s*a-a] 'to pile-DECL'

In (17b-c), the underlying stem-final /h/ deletes, due to an independent process that deletes stem-final /h/ when it is followed by a V-initial suffix, so that there is no intervening consonant between the trigger and the target vowel in the surface form. In such cases, where the stem ends with a vowel at the surface level, only the harmonic form is chosen.

In the current data, presented in Figure 5, the hiatus effect reported in previous research is largely confirmed. It was found that within each subcorpus, the disharmonic form was attested more frequently when there was at least one intervening consonant between the two vowels (VC(C)V) compared to the hiatus context (VV). However, it was not the case that all words in

the hiatus context were harmonic; 2 out of 153 words were suffixed with the disharmonic form, i.e. /nas-A/ → [na-ʌ] 'to be better-DECL (familiar)' (where the underlying stem-final /s/ deletes due to an independent process applied to the so-called 's-irregular' stems) and /tah-Ajo/ → [ta-ʌjo] 'to touch-DECL (honorific)'.

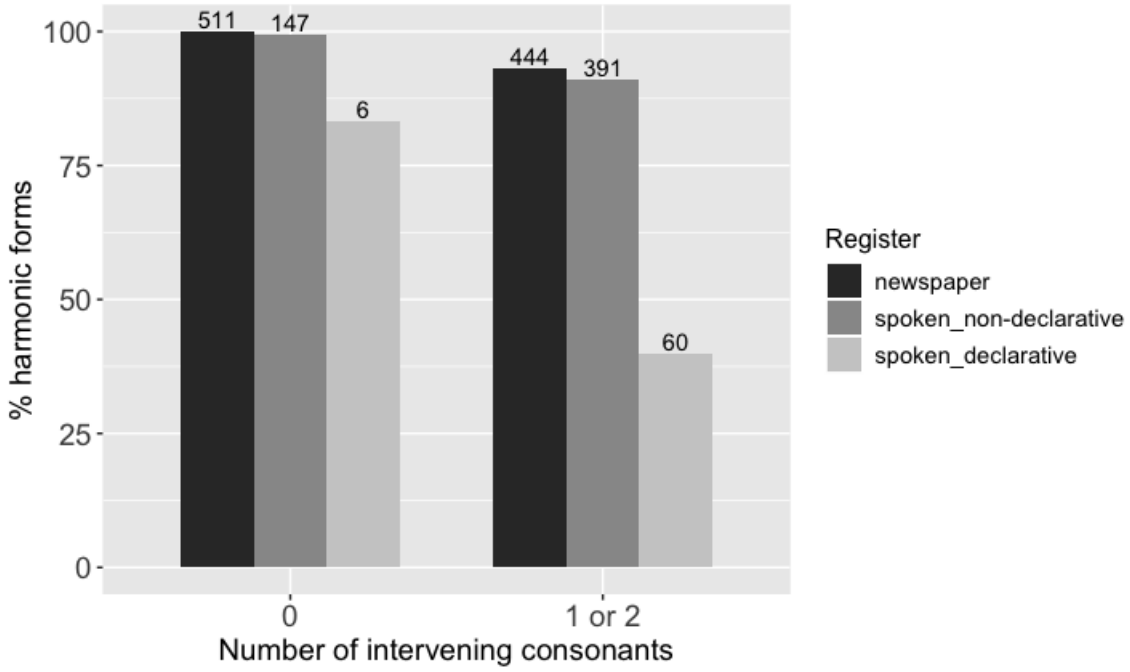

Figure 5: The effect of number of intervening consonants

It is noticeable that the hiatus effect is most evident in the Declarative data, as the difference in the harmony rate between VV and VC(C)V was the greatest in this subcorpus. This finding should be viewed with caution, though, as there are only six words that belong to the VV category.

If this near-categorical harmony in hiatus context is due to the adjacency of the trigger and the target vowel, we might expect a locality effect such that harmony is better observed when there is only one intervening consonant (e.g. /mak-A/ 'to block-DECL') compared to cases where two consonants intervene (e.g. /malk-A/ 'to be clear-DECL'). This prediction was not upheld in this study; there was only a small difference between VCV and VCCV conditions, and if anything, the tendency was in the opposite direction (see also Y. Kang and Ryu (2015)).

3.4.4 The special case of p-irregular stems

We turn now to the data involving p-irregular stems. Recall from section Section 3.1 and Section 3.2 above that p-irregular stems, in which the stem-final segment /p/ is realized as [w] when followed by V-initial suffixes, frequently combine with disharmonic suffix forms when the stem vowel is /a/ or /o/, as noted in previous studies (Hong, 2008; H. Kang, 2012). As shown in (18) (already presented above as part of (10)), the harmonic [a]-initial form is the minor variant for many p-irregular /a/- and /o/-stems.

(18) *P-irregular* /a/*-stems and* /o/*-stems frequently take disharmonic suffixes*

a. /pankap-A/ [pankaw-ʌ] (~[pankaw-a]) 'to be nice to see-DECL'

b. /welop-A/ [welow-ʌ] (~[welow-a]) 'to be lonely-DECL'

c. /kwelop-A/ [kwelow-ʌ] (~[kwelow-a]) 'to be in distress-DECL'

Among the p-irregular verbs, monosyllabic /o/-stems form an exception. As shown in (19), monosyllabic p-irregular /o/-stems combine only with [a]-initial suffix forms. (Monosyllabic p-irregular /a/-stems do not seem to exist in Korean.)

(19) *Monosyllabic p-irregular stems take harmonic suffixes*

a. /top-A/ [tow-a] 'to help-DECL'

b. /kop-A/ [kow-a] 'to be beautiful-DECL'

Simply put, the monosyllabic p-irregular stems exhibit ordinary vowel harmony.

Figure 6 shows an interaction of stem class (regular and p-irregular), stem vowel (/a/ and /o/) and stem length (monosyllabic and polysyllabic) across the three subcorpora. The lower rate of harmony for p-irregular stems is confirmed, and the exceptional behavior of monosyllabic p-irregular /o/-stems is also confirmed, although there are only two words in that category. The effect of p-irregular stem was robust in all three subcorpora.

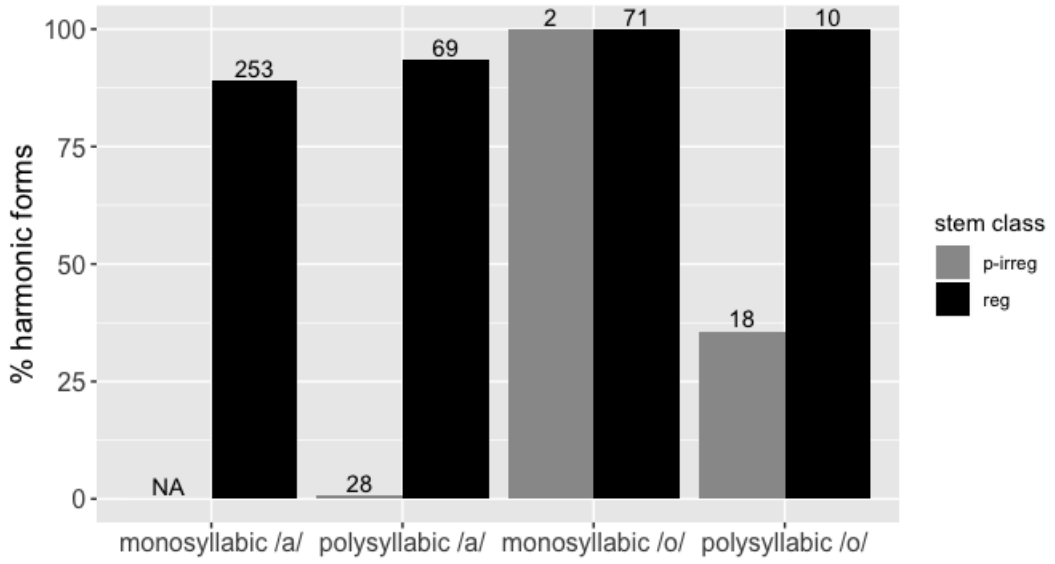

Figure 6: The effect of stem class by stem length and stem vowel

3.4.5 Is there a stem length effect?

It is observed in a subset of vowel harmony languages that root-initial or word-initial syllable controls the harmony (McCollum & Essegbey, 2020; Ringen & Heinämäki, 1999; Vaysman, 2009). Somewhat similarly, it is plausible in the current data that the vowel in initial syllable is a stronger trigger than those in other positions. If this is the case, monosyllabic stems are more likely to trigger harmony than polysyllabic stems, as the last stem vowel that triggers the harmony occupies the initial position in monosyllabic stems. One thing we can note in Figure 6 is that monosyllabic stems have a higher harmony rate than polysyllabic stems. However, I take this apparent stem length effect to be an artifact of the stem class effect. When only regular stems were considered, the harmony rate was 92.6% for monosyllabic stems and 97.0% for polysyllabic stems. Statistical testing also revealed that the effect of stem class was significant while that of stem length was not (see Section 3.5).[8]

[8]Such monosyllabicity effect was not found in the phonotactic data, either; in a phonotactic analysis in which only disyllabic stems were considered (not reported in this paper), where the trigger vowel always occurs in the initial position, they did not show a particularly stronger tendency to exhibit harmony.

3.4.6 Frequency of the stem

Figure 7 shows the relationship between log frequency of the stem and harmony rate.[9] For visual clarity, the data is jittered, i.e. a small random value is assigned to dots on the plot so that they are not plotted on top of each other. It is hard to make a generalization regarding the overall frequency effect on the harmony pattern, partly because the majority of words show categorical harmony. Nevertheless, it is possible to observe two weak tendencies. First, words with high stem frequency and those with low stem frequency tend to be either highly harmonic or highly disharmonic, rather than have an intermediate harmony rate. Words with intermediate stem frequency show more free variation. Second, among the words that have intermediate harmony rate, words with lower stem frequency tend to have lower rate of harmony. Although the tendency is rather weak, it may indicate that Korean vowel harmony is becoming less productive in the low frequency stems, in line with the claim that high frequency words tend to resist morphosyntactic changes (Bybee, 2001).

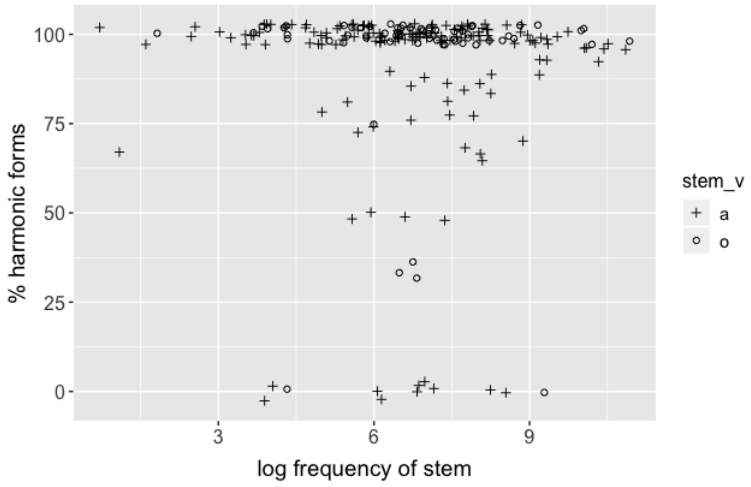

Figure 7: The effect of stem frequency

[9]The stem frequency was obtained from B.-M. Kang and Kim (2004).

3.4.7 Interim summary

Investigation of the variable pattern of harmony in /a/-stems and /o/-stems in the spoken corpus replicated findings of previous research (e.g. Hong, 2008; H. Kang, 2012) that /a/-stems discouraged harmony (i.e. trigger vowel effect) and that harmony rate is higher in the hiatus context (i.e. hiatus effect). In addition, it was also found that the two effects are most clearly observed in the Declarative spoken data. To summarize the findings presented thus far, the harmony is retreating in the /-A/ 'DECL' suffix, and within this suffix, the change is targeting particular phonological contexts while others are more resistant to rule breakdown.

3.5 Statistical testing

In order to test whether the generalizations described so far are statistically valid, I ran a logistic regression analysis with harmony rate as the dependent variable. More specifically, the statistical analysis aimed to test (i) whether the /-A/ 'DECL' significantly resisted harmony, (ii) whether harmony was more strictly observed in the newspaper corpus than in the spoken data, even when only suffixes other than the /-A/ 'DECL' were considered, and (iii) whether the phonological and morphological factors that regulate the variable pattern of harmony have a statistically significant effect on the harmony rate.

Independent variables are presented in (20), with a brief comment on the harmony patterns relevant to each predictor. All the predictor variables were dummy-coded (underlined level is the baseline). Log frequency of the stem was not included as a predictor, as it was hard to make a generalization about the effect of stem frequency on the harmony rate (see Section 3.4.6). [10]

[10] As an exploratory analysis, I also investigated potential effects of the place and manner feature of the intervening consonant. I did not present the results in the previous section, nor do I include the factors in the model. This is because they did not seem to have a substantial effect on the harmony rate (one previous study (H.-S. Kang, 2002) also did not find any noticeable effect of consonant type on the harmony pattern). Moreover, it was unclear how the type of consonant for VCCV words should be coded, making an assessment of its effect difficult.

(20) *Independent variables in the logistic regression analysis for alternation*

- SUBCORPUS (newspaper, Declarative (spoken), Non-declarative (spoken)): The harmony rate was highest in the newspaper data, and lowest in the Declarative data.
- STEM VOWEL (/a/ and /o/): /o/-stems showed a higher harmony rate.
- NUM C (the number of intervening consonants; 0, 1 or 2): Harmony rate was higher when the number of consonants was 0.
- SUBCORPUS × STEM VOWEL: The trigger vowel effect was active only in the Declarative data.
- SUBCORPUS × NUM C: The hiatus effect was stronger in the Declarative data.
- STEM TYPE (regular and p-irregular): P-irregular stems discouraged harmony.
- STEM LENGTH (monosyllabic and polysyllabic): Monosyllabic stems had a higher harmony rate.
- STEM TYPE × STEM LENGTH: Monosyllabic p-irregular stems still have a higher rate of harmony despite the fact that p-irregular stems tend to resist harmony.

I employed a logistic regression model using the *bayesglm()* function from the *arm* package (Gelman & Su, 2018) in R (R Core Team, 2018). I used *stepAIC()* in the MASS package (Venables & Ripley, 2002), using both forward and backward search in the stepwise model selection. The best model was identified as the full model minus SUBCORPUS × NUM C, indicating that the interaction between SUBCORPUS and NUM C did not significantly improve the model.[11]

The results of the statistical analysis are presented in (21). First, we see that the Non-declarative portion of the spoken data had a significantly lower harmony rate than the newspaper data and a significantly higher harmony rate than the Declarative data, when the stem vowel was /a/. The difference between the newspaper data and the Non-declarative data was rather small, as indicated by the small coefficient (see also Figure 3). Harmony was significantly encouraged for /o/-stems (in the Non-declarative data). Importantly, analysis confirmed that the trigger vowel effect was significantly stronger in the Declarative data, compared to the Non-declarative data,

[11] This result is rather unexpected, given that the discrepancy between VV and VC(C)V was the biggest in the Declarative data (see Figure 5). This seems to be due to lack of statistical power to assess this factor, as there were only a few words that are suffixed with the /-A/ 'DECL' and are in the hiatus context.

represented by the positive coefficient of the SUBCORPUS (Declarative) : STEM VOWEL (/o/). The significant interaction SUBCORPUS (newspaper) : STEM VOWEL (/o/) had a negative coefficient, suggesting that the trigger vowel effect was weaker in the newspaper data; this was not evident in Figure 4 and the effect size is not as large. The effect of NUM C was not significant (for the baseline of SUBCORPUS is Non-declarative in which harmony rate did not differ much between VV and VC(C)V words; see Figure 5).[12] The effect of STEM TYPE was significant, indicating that p-irregular stems significantly resisted harmony, while the effect of STEM LENGTH was non-significant. This result suggests that the apparent higher rate of harmony in monosyllabic stems was due to the fact that most p-irregular stems happened to be polysyllabic. Last, consistent with what we saw in Figure 6, the significant interaction between STEM TYPE and STEM LENGTH confirms that despite the tendency of p-irregular stems to discourage harmony, monosyllabic p-irregular stems still had a high harmony rate.

(21) *Logistic regression analysis for suffix alternation*

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	4.630	0.647	7.160	<.001
SUBCORPUS (newspaper)	1.537	0.741	2.074	< .05
SUBCORPUS (Declarative)	-4.687	0.512	-9.154	<.001
STEM VOWEL (/o/)	3.286	1.024	3.210	< .01
SUBCORPUS (newspaper) : STEM VOWEL (/o/)	-2.559	1.164	-2.198	<.05
SUBCORPUS (Declarative) : STEM VOWEL (/o/)	5.773	1.150	5.021	<.001
NUM C (0)	1.289	0.761	1.694	0.090
STEM TYPE (p-irregular)	-8.998	1.001	-8.990	< .001
STEM LENGTH (monosyllabic)	-0.864	0.591	-1.463	0.143
STEM TYPE (p-irregular) : STEM LENGTH (monosyllabic)	6.019	2.013	2.991	< .01

[12] I ran a separate regression model in which the baseline level for SUBCORPUS was Declarative, in order to test the effect of NUM C for the Declarative data. The effect of NUM C was not significant in this model, either. Given that the harmony rate in the Declarative data differed substantially between VV and VC(C)V contexts, at least descriptively (see Figure 5), the non-significant effect seems to be due to the rather small number of words in the Declarative subcorpus, not having enough statistical power.

To sum up the results of corpus study for suffix alternation, the newspaper corpus exhibited a significantly higher harmony rate than spoken data, suggesting that the vowel harmony process is observed more strictly in the more formal and conservative register of Korean. Within the spoken corpus, words suffixed with the /-A/ 'DECL' had a significantly weaker harmony than words with other suffixes; it seems that this suffix is leading the breakdown of the harmony process. Further, the trigger vowel effect whereby the suffix vowel was more likely to be harmonic for /o/-stems was only observed in the Declarative portion of the spoken data. Although the effect of the number of consonants or its interaction with the subcorpus factor was not statistically significant, the pattern was evident that the hiatus effect was stronger in the Declarative data. In the next section, I assess the effects of these two phonological factors on the tendency for two vowels to harmonize in the phonotactics of Korean, in addition to investigating whether vowels in the lexicon show a tendency to harmonize at all.

4 A comparison between alternation and phonotactics

The main purpose of phonotactic analysis was twofold. First, I sought to test whether Korean lexicon is subject to gradient vowel co-occurrence restrictions that reflect the harmony pattern. Second, I tested the possibility that the effects that condition the variable harmony pattern in the /-A/ 'DECL', i.e. the effect of the trigger vowel and the number of intervening consonants, originate from the lexicon.

To analyze phonotactics in the lexicon, two subsets of the frequency data of Korean (B.-M. Kang & Kim, 2004) were used, i.e. the list of content morphemes (henceforth, All Stems data) and that of verbal and adjectival stems (henceforth, Verb Stems data). From the two lists, I only extracted stems with a frequency of 10 or higher. All the vowel sequences V_1V_2 whose V_2 was either [a] or [ʌ] were identified, along with the information about the number of consonants intervening between the two vowels. By considering only cases where V_2 was either [a] or [ʌ], which are the two allomorph vowels of the suffix alternation, the phonotactic analysis was made

as parallel as possible to the alternation. This resulted in a total of 42,515 stems in All Stems data and 2,355 stems in Verb Stems data.

4.1 Gradient harmony in phonotactics

As a first step of comparing alternation and phonotactic patterns, I first examined how V_2 is realized depending on the [ATR]/[RTR] specification of V_1 in the two phonological generalizations (for alternation, stem vowel and suffix vowel are considered V_1 and V_2, respectively). In the following figures, the y-axis represents the proportion of [a] forms in V_2, rather than the harmony rate. The numbers at the top of each bar indicate the number of V_1V_2 sequences that belong to that category. The harmony rates for alternation in V_1=[RTR] cases should be the same as those of /a/-stems and /o/-stems presented in Section 3.4.

Figure 8 compares the proportion of [a] in V_2 in the two sets of alternation data (Declarative and Non-declarative) with the two sets of phonotactic data (All Stems and Verb Stems). It can be seen that for both sets of phonotactic data, the proportion of [a] in V_2 is higher when preceded by [RTR] vowels compared to when preceded by [ATR] vowels, suggesting a tendency towards harmony. When V_1 was a [RTR] vowel, the proportion of [a] in V_2 was 78.6% for Verb Stems and 64.9% for All Stems data; these rates are higher than the Declarative but lower than the Non-declarative subcorpus. We can thus say that phonotactic harmony is not as strong as the canonical pattern of harmony that is observed in Non-declarative suffixes. We can also say that the low harmony rate of the /-A/ 'DECL' suffix cannot be attributed to lack of harmony in the lexicon, since stem phonotactics shows at least a tendency towards harmony.

As we might expect, alternation and phonotactics exhibit a huge discrepancy in terms of the proportion of [a] following an [ATR] vowel. As presented in (9) (Section 3.1) and (11) (Section 3.2), when the stem vowel is [ATR], the suffix vowel is rarely [a]. In contrast, Figure 8 shows that an [ATR] V_1 in phonotactics is followed by [a] quite frequently. Given than [RTR] vowels are triggers of the harmony, it would not be too surprising to see a moderately high percentage of [a] after [ATR] vowels in the phonotactic data, because there is no trigger to motivate the har-

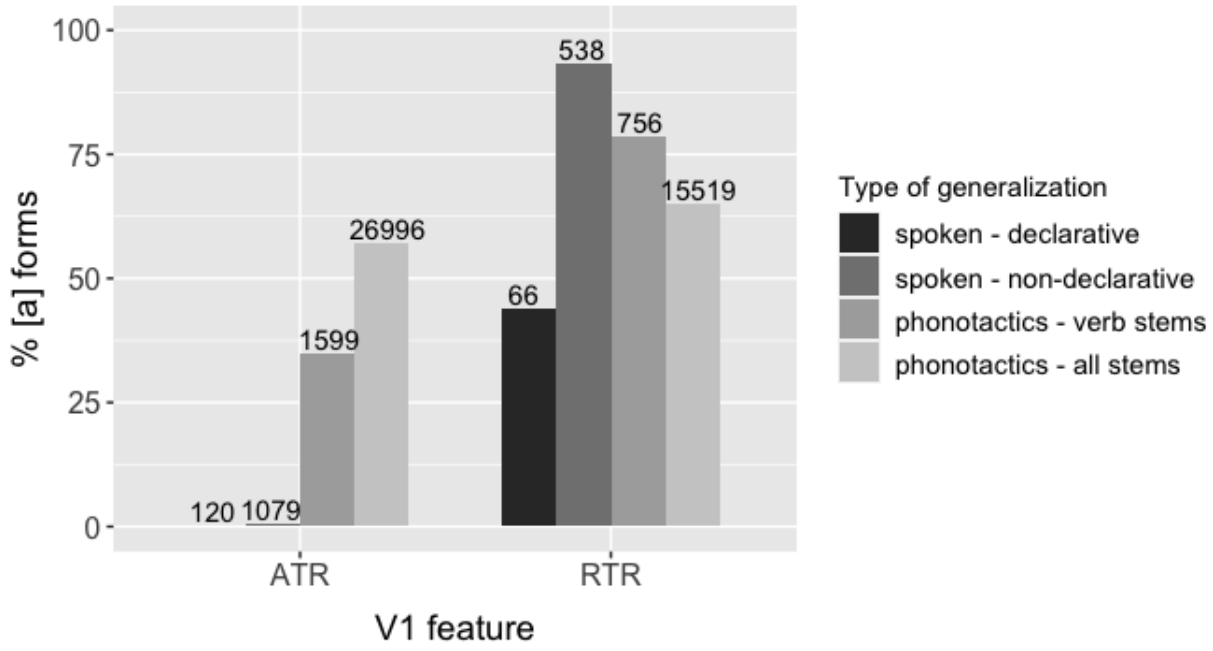

Figure 8: The percentage of [a] forms in V_2

mony in this context. It still holds true, however, that the categorical harmony in which [ATR] stem vowel is followed by an [ATR] suffix vowel lacks phonotactic support. Of the two sets of phonotactic data, it seems that Verb Stems data more closely match the harmony pattern in the alternation, compared to All Stems data.

To summarize, for the Non-declarative subcorpus, there is a huge difference in the proportion of [a] between V_1=[ATR] and V_1=[RTR], in the harmonic direction; for Verb Stems, there is still a substantial difference in the same direction (but not as notable as that of the Non-declarative data); and for All Stems, there is a small difference in the same direction.

To assess whether the observed effects are statistically significant, a logistic regression analysis was implemented, with the proportion of [a] in V_2 as the dependent variable. The model included two independent variables, both of which were dummy-coded. The first is DATA with four levels, Declarative, Non-declarative (baseline), All Stems and Verb Stems. The other independent variable was V_1 FEATURE with two levels, ATR (baseline) and RTR. The interaction terms between DATA and V_1 FEATURE were also included, which were expected to be significant

if the effect of V_1 FEATURE on the proportion of [a] forms in V_2 was significantly stronger (i.e. the tendency to harmonize is stronger) in Non-declarative compared to other levels.

The results of the analysis are provided in (22). As expected, the effect of V_1 FEATURE was significant, indicating that the proportion of [a] in V_2 is higher in V_1=[RTR] context compared to V_1=[ATR], when the data is Non-declarative.[13] Importantly, the interaction between DATA (All Stems) and V_1 FEATURE (RTR) was significant, indicating that a [RTR] V_1 enhanced the proportion of [a] in V_2 to a greater degree in the Non-declarative data relative to All Stems. Similarly, the interaction between DATA (Verb Stems) and V_1 FEATURE (RTR) was also significant, suggesting that the effect of V_1 FEATURE was stronger in the Non-Declarative data compared to the Verb Stems data.

(22) *Logistic regression analysis: harmony in alternation vs. phonotactics*

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-5.476	0.472	-11.591	< .001
DATA (All Stems)	5.764	0.473	12.198	< .001
DATA (Verb Stems)	4.852	0.475	10.208	< .001
DATA (Declarative)	-8.091	48.878	-0.166	0.869
V_1 FEATURE (RTR)	8.094	0.053	16.108	< .001
DATA (All Stems) : V_1 FEATURE (RTR)	-7.768	0.503	-15.444	< .001
DATA (Verb Stems) : V_1 FEATURE (RTR)	-6.172	0.513	-12.031	< .001
DATA (Declarative) : V_1 FEATURE (RTR)	5.226	48.879	0.107	0.915

In sum, a comparison between alternation and stem phonotactics revealed that although tendency to harmonize is observed in both, it is more robust in alternation, specifically in suffixes other than the /-A/ 'DECL'. Therefore, we can conclude that the harmony shown in alter-

[13] To test whether the effect of V_1 FEATURE is significant in the two sets of phonotactic data, I established two separate models, not presented here, one of which had Verb Stems as the baseline level and the other having All Stems as the baseline. The effect of V_1 FEATURE was significant in both models, suggesting that both phonotactic data indeed exhibited harmony.

nation is partly supported by phonotactics in the lexicon. In the following section, I investigate whether the trigger vowel effect and the hiatus effect are present in the lexicon by comparing the two sets of phonotactic data against the /-A/ 'DECL' data, in which the two effects emerged most robustly.

4.2 No trigger vowel effect or hiatus effect in stems

Recall from Section 3.4.2 that /o/ is a stronger trigger than /a/ in alternation of the /-A/ 'DECL' suffix (also seen as black bars in Figure 9). Such an effect of trigger vowel is not observed in the phonotactic data, as can be seen in Figure 9. The proportion of [a] in V_2 is little different for [a] and [o] trigger vowels. Therefore, alternation and phonotactics exhibit an asymmetrical effect of the trigger vowel.

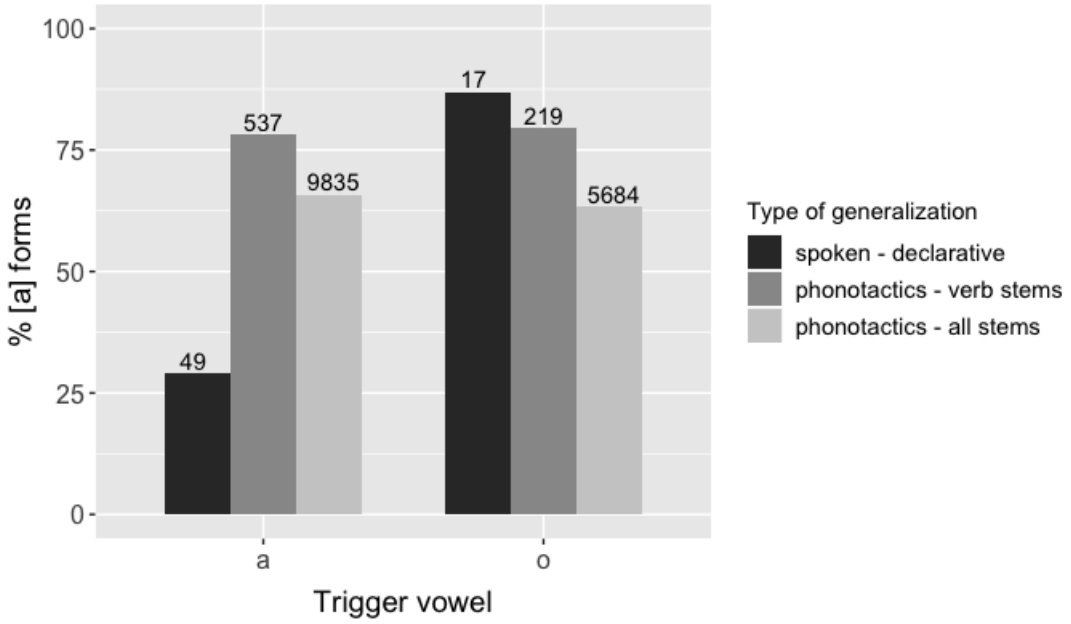

Figure 9: The trigger vowel effect in alternation and phonotactics

Figure 10 compares phonotactics against alternation with regard to the effect of the number of intervening consonants between two vowels. Recall from Section 3.4.3 (see also the black bars in Figure 10) that in the /-A/ 'DECL' suffix, the harmony rate was much higher when the suffix vowel immediately follows the stem vowel, compared to when they are separated by one

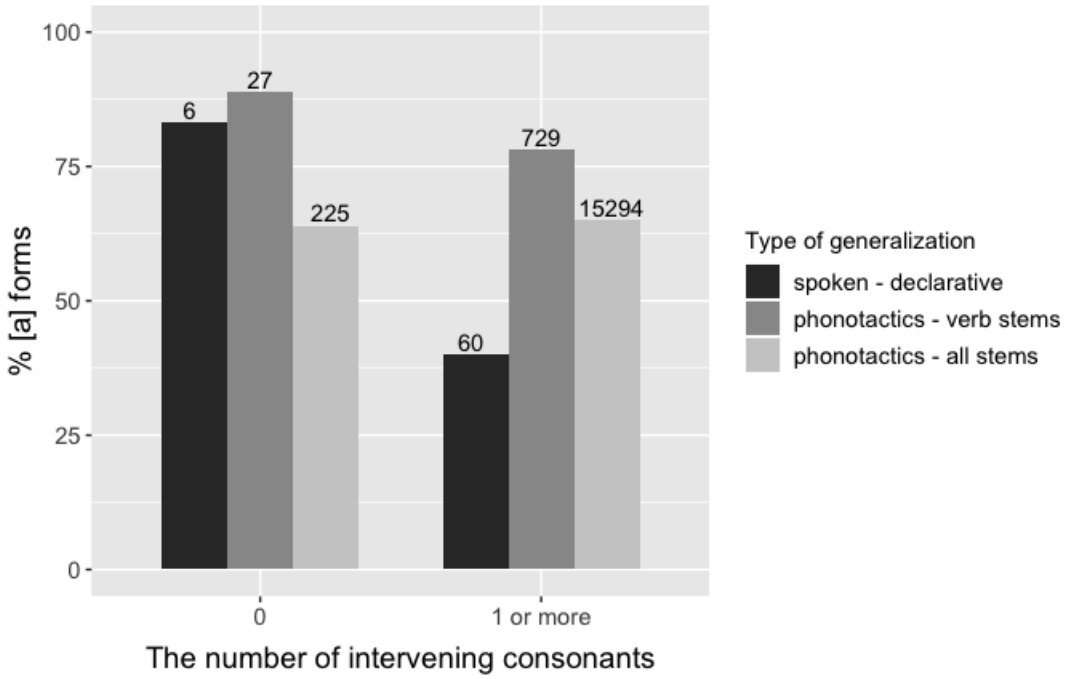

Figure 10: The hiatus effect in alternation and phonotactics

or more consonants. Such hiatus effect did not emerge in the All Stems data as the harmony rate was very similar in the hiatus and the non-hiatus contexts and emerged only weakly in the Verb Stems data, as the harmony rate was slightly higher for the hiatus context. Note also that for neither of the phonotactic datasets, the harmony rate was near 100% in the hiatus condition. This is in contrast with pattern in the newspaper corpus and the Non-declarative portion of the spoken corpus, in which two adjacent vowels showed a near-categorical harmony. Therefore, the lexicon hardly matches the alternation pattern in terms of the hiatus effect.

A logistic regression analysis was carried out to test whether the trigger vowel effect and the hiatus effect were active only in the alternation data. The dependent variable was the proportion of [a] in V_2. Independent variables (dummy-coded) were the following (with the reference level of each factor underlined): DATA (Declarative, All Stems and Verb Stems), V_1 (/a/, /o/), NUM C (i.e. the number of intervening consonants; 0, 1 or more), and two interactions DATA : V_1 and DATA : NUM C. If the trigger vowel modulates the proportion of [a] in V_2 only in the Declarative data, or more strongly so, DATA : V_1 will test as significant. The same holds true for

the hiatus effect and DATA : NUM C.

The table in (23) presents the results. First, the significant effects of DATA showed that both All Stems and Verb Stems data had a higher rate of harmony than the alternation data, in /a/-stems and in non-hiatus contexts. The effect of V_1 was significant, while that of NUM C missed significance. The two interaction terms DATA (All Stems) : V_1 (/o/) and DATA (Verb Stems) : V_1 (/o/) were significant, indicating that /o/ in V_1 encouraged harmony only in the Declarative data but not in the two phonotactic datasets. The interaction of DATA (All Stems) : NUM C (0) and of DATA (Verb Stems) : NUM C (0) were non-significant, despite the pattern seen in Figure 10 in which the difference between hiatus and non-hiatus context was the greatest in the Declarative data. I speculate that it was hard to statistically assess the interaction between DATA and NUM C, as was the case in Section 3.5, because there were only a few words (N=6) in the Declarative subcorpus that belong to the NUM C = 0 category.

(23) *Logistic regression analysis: the trigger vowel effect and the hiatus effect in alternation vs. phonotactics*

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-1.111	0.345	-3.223	< .01
DATA (All Stems)	1.768	0.345	5.121	< .001
DATA (Verb Stems)	2.370	0.360	6.577	< .001
V_1 (/o/)	2.883	0.801	3.597	< .001
NUM C (0)	2.256	1.189	1.898	0.058
DATA (All Stems) : V_1 (/o/)	-2.996	0.802	-3.734	< .001
DATA (Verb Stems) : V_1 (/o/)	-2.821	0.825	-3.418	< .001
DATA (All Stems) : NUM C (0)	-2.286	1.197	-1.909	0.056
DATA (Verb Stems) : NUM C (0)	-1.461	1.341	-1.090	0.278

In sum, the trigger vowel effect and the hiatus effect that emerged in the Declarative subcorpus was barely observed in the lexicon. This finding suggests that the nuanced pattern of

variation in harmony observed in the suffix alternation is not well supported by stem phonotactics in the lexicon. Therefore, I reject the hypothesis that the two effects that began to be attested in the spoken data were motivated by phonotactic generalizations.

5 Productivity tests in previous research

In this section, I review productivity tests reported in H. Kang (2012) and H. Jang (2017), which tested native speakers' intuition on the harmony pattern, in light of the findings of this study. The results of these tests indicate that the harmony pattern is not fully reproduced in the wug tests. In the production experiment in H. Kang (2012), two striking results were found. First, when the /-A/ 'DECL' was employed, the proportion of harmonic forms was only 32.0% for nonce /a/- and /o/- stems, as compared to 76.7% for real stems. This result does not in itself suggest that the harmony is not internalized in the speakers' grammar, because they always produced the harmonic [ʌ]-forms for [ATR] stem vowels such that proportion of [a] response did differ between [ATR] and [RTR] stem vowels (i.e. 32.0% vs. 0%). However, the rate of harmony in the wug test for [RTR] stem vowels is far lower than for real words, suggesting that the harmony has a limited productivity.

Second, in the same study, unlike in real stems where trigger vowel effect emerged, nonce /a/-stems and /o/-stems were very similar in their probability of taking a harmonic suffix. Both had a harmony rate of approximately 32%, although this result might not be as reliable due to the small number of stimuli (fifteen /a/-stems and five /o/-stems were used). We have seen that the trigger vowel effect was very weak in phonotactics, which leads to the conjecture that the lack of difference between /a/ and /o/ as trigger vowel in the lexicon may be one factor that accounts for the absence of trigger vowel effect in nonce stems.

Another wug test, presented in H. Jang (2017), also demonstrates a low productivity of the harmony. In this experiment, as was the case in Kang's experiment, speakers frequently produced disharmonic suffixes for /a/- and /o/-stems. What is striking is that a few speakers even

volunteered the disharmonic [a]-form for [ATR] stem vowels, despite the exceptionless harmony pattern for [ATR] stem vowels in real words. Once we consider that the [ATR] harmony is maintained only weakly in the lexicon, however, such aberrant behavior of the speakers may not be very surprising; lack of phonotactic grounds for a perfect [ATR] harmony might have led some speakers to disobey it.

In sum, existing wug tests on the harmony pattern directly demonstrate that it is not very productive. I argue, based on the findings of Section 4, that the low productivity of suffix alternation might be ascribed to the lack of robust phonotactic support.

6 Discussion and Conclusions

The current study investigated the variable pattern of vowel harmony in Contemporary Korean in the newspaper corpus and in the spoken corpus and examined gradient vowel co-occurrence restrictions in the lexicon that might reflect the harmony pattern. The table in (24) summarizes the results of this study for cases in which the trigger vowel (or, V_1 for phonotactics) is [a, o]. The overall harmony rate was highest in the suffix alternation in the newspaper corpus (Section 3.1), which is presumably a conservative register in Korean. The harmony rate was slightly lower in the Non-declarative portion of the spoken data, but harmony still applied quite regularly; in comparison, the Declarative portion of the spoken data had a significantly low rate of harmony, suggesting that the harmony has begun to retreat in the /-A/ 'DECL' suffix (Section 3.2). In the wug test reported in H. Kang (2012), as reviewed in Section 5, the harmony rate was even lower. The continuum of harmony rate conditioned by register, suffix type and the real/nonce status of the stem adds another piece of evidence to the existing literature on decay of the harmony. The harmony rates in phonotactics were intermediate, falling between harmony rates of highly-harmonic alternation data and those of weak-harmony alternation data (Section 4). The phonotactic restrictions in the lexicon, therefore, support the standard pattern of harmony of the suffix alternation only moderately. The findings suggest that loss of productivity in alterna-

tion may be attributed to the fact that stem phonotactics matches the alternation pattern only weakly.

Moreover, this study also showed that the effects of the two phonological factors that were reported to regulate the harmony pattern in previous research (M.-H. Cho, 1994; Hong, 2008; H. Kang, 2012; Kim-Renaud, 1976), i.e. the trigger vowel effect and the hiatus effect, were most robust in the /-A/ 'DECL' suffix (Section 3.4). The trigger vowel effect was exclusively attested in the Declarative data. The hiatus effect was strongest in the Declarative data, and emerged only weakly in the newspaper corpus and the Non-declarative data. These two effects were rarely found in the two phonotactic datasets, aside from the fact that a weak hiatus effect was observed in the Verb Stems data (Section 4). This finding suggests that the phonotactics does not support the nuanced pattern of variation in the alternation; the two characteristics of the breakdown of harmony, i.e. it has begun in the /a/-stems and in non-hiatus contexts, are not phonotactically motivated. The effect of stem type (specifically, the tendency of p-irregular stems to resist harmony) was robust in all alternation datasets, and the effect of intervening [w] between two vowels in the lexicon remains to be tested in future works.

(24) *Summary of the harmony pattern in alternation and phonotactics*

	newspaper	spoken		wug test (Decl.)	phonotactics	
		Non-decl.	Decl.	H. Kang (2012)	Verb Stems	All Stems
harmony	strong (97%)	strong (93%)	weak (44%)	weak (32%)	intermediate (79%)	intermediate (65%)
trigger vowel effect	no	no	yes	no	no	no
hiatus effect	weak	weak	strong	not tested	weak	no
stem type effect	yes	yes	yes	N/A	(future work)	(future work)

The findings of the present study are in line with the claim of Chong (2019) that a derived

environment rule is productive to the extent that it is supported by phonotactics; a weak phonotactic support for the alternation might have led to the gradual loss of productivity of Korean vowel harmony. To recap Chong (2019), in the known cases of derived environment effects, the reported mismatches between phonotactics and alternations are only superficial, as exemplified by Korean palatalization mentioned in Section 1.1. He also argues that when the phonotatics-alternation mismatch is real, as in Turkish velar deletion, the alternation actually occurs in limited contexts, evidenced by the finding that Turkish velar deletion applies productively only in polysyllabic nouns. Similar to the case of Turkish velar deletion, Korean vowel harmony occurs in very restricted contexts, as discussed in detail in Section 2. One apparent cause of this limited productivity is that the phonotactic generalizations only weakly match alternation patterns; the tendency towards harmony was not completely absent in the lexicon, but it emerged as a gradient constraint such that its effect was not as strong as that of the alternation. Such feeble phonotactic support may also account for its ongoing loss of productivity as shown in this study and in existing wug tests that demonstrate a lower harmony rate in nonce stems compared to real stems (H. Jang, 2017; H. Kang, 2012, see Section 5).

We observed that breakdown of the harmony rule was more obvious in some contexts than in others. First, it was found that the harmony was weakest in the /-A/ 'DECL' suffix, which was attested only in the spoken corpus. Within the spoken corpus, the /-A/ 'DECL' portion represents data that is exclusively casual, whereas for many of the other suffixes included in the analysis (see (8)) the speech style could be a mix of casual and formal one. The /-A/ 'DECL' is used in non-honorific utterances, and only in vernacular speech in which two intimate speakers converse informally; most of the other suffixes are exclusively used in the honorifics or can be used in both honorifics and non-honorifics as they do not encode politeness of the utterance (Seo, 1984; Sohn, 1999). Given that in typical stages of rule death, the breakdown of a phonological process begins in the most vernacular register of the language (Labov, 1966, 1972; Maclagan, 2000; Rohena-Madrazo et al., 2006), Korean vowel harmony seems to be at the beginning stage of rule breakdown in which the harmony rate was highest in the newspaper data

(which is plausibly thought to be the most formal and conservative register) and lowest in the spoken data with the /-A/ 'DECL'. It was also this suffix that leads the trigger vowel effect and the hiatus effect, the two major factors that condition the variation.

As an alternative account for why the harmony is weakest in the /-A/ 'DECL' suffix, H. Kang (2012) explains its lower harmony rate by appealing to the fact that it is the only one among the harmonizing suffixes (see (8)) in which the target vowel occurs sentence-finally (in other suffixes, the target vowel occurs either word-medially, or clause-finally but not sentence-finally.) He argues that the target vowel in the /-A/ 'DECL' occupies the final position which is prosodically salient due to final lengthening in Korean, and is thus governed by positional faithfulness constraint (Beckman, 1997) to be faithfully realized as [ʌ] (assuming that /ʌ/ is the underlying form). The alternative hypothesis I am suggesting here is that this is simply part of the general pattern that lower harmony rates accompany more vernacular styles; the suffix type effect could be actually in large part a register effect.

Within the /-A/ 'DECL' portion, it was in the /a/-stems that the harmony was retreating. It is plausible to imagine that the difference between /a/-stems and /o/-stems in their strength as a trigger is motivated by phonotactic generalizations. This hypothesis was not supported, because phonotactic data did not exhibit any difference between [a] and [o] in triggering harmony. Then, why does the breakdown begin in /a/-stems, while /o/-stems still maintain the harmony robustly? It is not easy to understand this trigger vowel effect. If anything, it is expected that [o-a] should be relatively dispreferred, and [a-a] should be relatively more preferred for the following reasons: (i) given the vowel system of Korean where [o] is more similar to [ʌ] than to [a] (see (3)), the process by which /o-ʌ/ surfaces as [o-a] almost seems like a disharmony, and (ii) /a-ʌ/ surfacing as [a-a], which is assimilatory, should be favored than /a-ʌ/ surfacing as [a-ʌ].

To understand the stronger trigger vowel effect of /o/-stems, H. Kang (2012) conducted a perception experiment in which the listeners were asked to identify [a] and [o] when the following vowel was either the harmonic [a] or the disharmonic [ʌ]. In adverse listening conditions (e.g. masking), listeners were better at identifying [o] when the following vowel was [a] com-

pared to when it was [ʌ]. The result suggests that [o] prefers to be followed by the harmonic [a] in order to enhance its perceptual accuracy, which might explain why [o] is a strong trigger. In contrast, [a] was accurately identified regardless of whether the following vowel was [a] or [ʌ]; thus, it can take a disharmonic form without compromising its perceptual accuracy. Although these results are not well explained with phonetic motivation, they are in line with the observation that /o/-stems are more likely to take the harmonic suffix form compared to /a/-stems.

To conclude, the current study has documented the beginning stage of the decay of Korean vowel harmony. The breakdown of harmony has begun in casual speech, in which we find the harmony is weakest in certain phonological contexts within this informal register. Based on the finding that the harmony pattern in suffix alternation is supported by the lexicon only moderately, I suggest that the loss of productivity in alternation is attributed to the fact that stem phonotactics only weakly matches the alternation pattern.

References

Ahn, S.-C. (1985). *The interplay of phonology and morphology in Korean* (Doctoral dissertation). University of Illinois at Urbana-Champaign.

Baek, E.-J. (1999). *Hankwuke yeksaumwunlon [Korean historical phonology]*. Seoul: Pakijeong.

Beckman, J. N. (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology*, *14*, 1–46. doi: 10.1017/S0952675797003308

Bybee, J. (2001). Frequency effects on French liaison. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 337–359). Amsterdam: John Benjamins.

Chae, W. (2003). *Hankwukeuy uysengewa uythaye [Korean onomatopoeia and mimetic words]*. Seoul: Seoul National University.

Cho, M.-H. (1994). *Vowel harmony in Korean: A grounded phonology approach* (Doctoral Dissertation). Indiana University.

Cho, S.-M. (2001). Choyceksengiloney uyhan moumcohwauy pyenhwa pwunsek [An

optimality-theoretic analysis on the change of the vowel harmony in Korean]. *Studies in Phonetics, Phonology and Morphology*, *7*(1), 191–213.

Cho, Y.-m. Y. (2016). Korean phonetics and phonology. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. doi: 10.1093/acrefore/9780199384655.013.176

Chong, A. J. (2019). Exceptionality and derived environment effects: A comparison of Korean and Turkish. *Phonology*, *36*(4), 543–572. doi: 10.1017/S0952675719000289

Chung, C.-W. (2000). An Optimality-Theoretic account of vowel harmony in Korean ideophones. *Studies in Phonetics, Phonology and Morphology*, *6*(2), 431–450.

Dresher, B. E. (1985). *Old English and the theory of phonology*. London: Routledge.

Foulkes, P., & Vihman, M. (2015). First language acquisition and phonological change. In P. Honeybone & J. Salmons (Eds.), *The oxford handbook of historical phonology* (pp. 1–32). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199232819.013.001

Gelman, A., & Su, Y.-S. (2018). arm: Data analysis using regression and multilevel/hierarchical models [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=arm` (R package version 1.10-1)

Han, E. (2009). Vowel harmony in hiatus contexts in Korean verbal morphology. *Studies in Phonetics, Phonology and Morphology*, *15*(2), 341–360.

Hock, H. H. (1991). *Principles of historical linguistics* (2nd ed.). The Hague: Mouton de Gruyter.

Hong, S.-H. (2008). Variation and exceptions in the vowel harmony of Korean suffixes. *The Journal of Studies in Language*, *24*, 405–428.

Hong, S.-H. (2010). Gradient vowel cooccurrence restrictions in monomorphemic native Korean roots. *Studies in Phonetics, Phonology and Morphology*, *16*(2), 279–295.

Jang, H. (2017). Decay of ungrounded rules: Middle Korean vowel harmony in contemporary Korean. In M. Y. Erlewine (Ed.), *Proceedings of GLOW in Asia XI, volume 2* (pp. 87–101).

Jang, Y. (1993). Kwuke moumchwuiwa moumcohwa pwungkoyey tayhaye [The vowel shifts and the breakdown of vowel harmony in Korean]. *Journal of Dong-ak Language and Litera-*

ture, 28, 45–65.

Jun, J. (2018). Morpho-phonological processes in Korean. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics.* New York: Oxford University Press.

Kang, B.-M., & Kim, H.-G. (2004). *Hankwuke hyengtayso mich ehwi sayong pintouy pwunsek2 [Frequency analysis of Korean morpheme and word usage2].* Seoul: Institute of Korean Culture, Korea University.

Kang, H. (2012). *Diachrony in synchrony: Korean vowel harmony in verbal conjugation* (Doctoral dissertation). Stony Brook University.

Kang, H. (2016). Hankwuke yongenhwalyongeyse poinun moumcohwauy pyeniwa pyenhwa: Hyencangsiwa silceysi yenkwu [Variation and change of Korean vowel harmony in verbal conjugation: An apparent time and real time study]. *Eoneohak, 76*, 27–56. doi: 10.17290/jlsk.2016..76.27

Kang, H.-S. (2002). A study of vowel harmony in the Modern Seoul Dialects. *The Korean Language and Literature, 26*, 1–27.

Kang, Y., & Ryu, N.-Y. (2015). Variation in Korean vowel harmony and the emergent locality effect. In *The joint meeting of the ICKL (International Circle of Korean Linguistics) and the Harvard-ISOKL (International Symposium on Korean Linguistics).* Chicago, Illinois, Jul. 24-26.

Kim, C.-W. (1973). Adjustment rules in phonology. In M. Kenstowicz & C. Kisseberth (Eds.), *Issues in phonological theory* (pp. 130–144). The Hague: Mouton.

Kim, J. S. (2007). Vowel harmony in Korean verbs. *Studies in Phonetics, Phonology and Morphology, 14*(3), 433–451.

Kim, W. (1971). *Kwukeumwuncheykyeyuy yenkwu [Studies on the phonological system of Korean].* Seoul: Ilcokak.

Kim, W. (1978). Reflections on the vowel system and vowel harmony of Korean. *Language Research, 14*(2), 127–139.

Kim, Y.-S. (1984). *Aspects of Korean morphology* (Doctoral dissertation). University of Texas,

Austin.

Kim-Renaud, Y.-K. (1976). Semantic features in phonology: Evidence from vowel harmony in Korean. In S. A. Mufwene, C. A. Walker, & S. B. Steever (Eds.), *Papers from the Twelfth Regional Meeting of the Chicago Linguistic Society* (pp. 397–412).

King, R. (1969). *Historical linguistics and generative grammar*. Englewood Cliffs, NJ: Prentice-Hall.

Kiparsky, P. (1971). Historical linguistics. In W. O. Dingwall (Ed.), *A survey of linguistic science.* College Park, MD: University of Maryland Press.

Kiparsky, P. (1973). Abstractness, opacity, and global rules. In O. Fujimura (Ed.), *Three dimensions in linguistic theory* (pp. 57–86). Tokyo: TEC.

Kiparsky, P. (1993). Blocking in non-derived environments. In E. M. Kaisse & S. Hargus (Eds.), *Studies in lexical phonology* (pp. 277–313). San Diego: Academic Press.

Labov, W. (1966). *The social stratification of English in New York City*. Wasington, D.C.: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lee, B.-G. (1985). Moumcohwauy thukseng [Some tangible properties of vowel harmony]. *Inmunnonchong*, *15*, 3–36.

Lee, J.-S. (1992). *Phonology and sound symbolism of Korean ideophones* (Doctoral dissertation). Indiana University.

Lee, K.-M. (1998). *Kwukesakaysel [A survey of the history of the Korean language]*. Seoul: Thaehaksa.

Lee, K.-M., & Ramsey, S. R. (2011). *A history of the Korean language*. Cambridge: Cambridge University Press.

Lee, S. N. (1947). Moumcohwayenkwu [Studies in vowel harmony]. *Cintanhakpo*, *16*, 1–109.

Lightfoot, D. (1999). *The development of language: Acquisition, change, and evolution*. Oxford: Blackwell Publishers.

Łubowicz, A. (2002). Derived environment effects in Optimality Theory. *Lingua*, *112*, 243–280.

Maclagan, M. A. (2000). Where are we going in our language? New Zealand English today. *NZ Journal of Speech-Language Therapy*, *53*, 14–20.

Mascaro, J. (1976). *Catalan phonology and the phonological cycle* (Doctoral dissertation). Massachusetts Institute of Technology.

McCarthy, J. J. (1983). Phonological features and morphological structure. In J. F. Richardson, M. Mitchell, & A. Chukerman (Eds.), *Papers from the Parasessions on the Interplay of Phonology, Morphology, and Syntax* (pp. 135–161). Chicago: Chicago Linguistic Society.

McCarthy, J. J. (2003). Comparative markedness. *Theoretical Linguistics*, *29*(1-2), 1–51. doi: 10.1515/thli.29.1-2.1

McCollum, A., & Essegbey, J. (2020). Initial prominence and progressive vowel harmony in Tutrugbu. *Phonological Data and Analysis*, *2*(3), 1–37. doi: 10.3765/pda.v2art3.14

O'Bryan, M. (1974). Opacity and rule loss. *Studies in the Linguistics Sciences*, *4*(2), 148–160.

Park, N. (2020). *Hankwuke umsopayyelceyyakuy thongkyeycek haksupkwa cekhyengseng phantan [Statistical learning of Korean phonotactics and well-formedness judgment]* (Doctoral Dissertation). Seoul National University.

Paster, M. (2013). Rethinking the 'duplication problem'. *Lingua*, *126*, 78–91. doi: 10.1016/j.lingua.2012.11.015

Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. New Brunswick, NJ: Rutgers Center for Cognitive Science, Rutgers University.

Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*.

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Ringen, C. O., & Heinämäki, O. (1999). Variation in Finnish vowel harmony: An OT account. *Natural Language and Linguistic Theory*, *17*, 303–337.

Rohena-Madrazo, M., Simonet, M., & Paz, M. (2006). The vernacular in the laboratory: The case of coda /r/ in Puerto Rico. In *Paper presented at the 35th annual conference on New Ways*

of Analyzing Variation (NWAV 35). Columbus, Ohio, Nov. 9-12.

Sen, R. (2016). Examining the life cycle of phonological processes: considerations for historical research. *Papers in Historical Phonology, 1*, 5–36. doi: 10.2218/pihph.1.2016.1691

Seo, J.-S. (1984). *A study on honorifics: The current honorific system and its problems [Contaypepuy yenkwu: Hyenhayng taywupepuy cheykyeywa mwunceycem]*. Seoul: Hansin Munhwasa.

Shin, J. (2015). Vowels and consonants. In L. Brown & J. Yeon (Eds.), *The handbook of Korean linguistics* (pp. 3–21). Oxford: Wiley Blackwell.

Shin, J., Kiaer, J., & Cha, J. (2013). *The sounds of Korean*. Cambridge, UK: Cambridge University Press.

Sohn, H.-M. (1999). *The Korean language*. Cambridge, UK: Cambridge University Press.

Vaysman, O. (2009). *Segmental alternations and metrical theory* (Doctoral dissertation). Massachusetts Institute of Technology.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer. Retrieved from `http://www.stats.ox.ac.uk/pub/MASS4` (ISBN 0-387-95457-0)