The Sino-Korean influence on Middle Korean vowel harmony: A usage-based perspective

Eric D. de Roulet

California State University, Fullerton

Abstract

While Early Middle Korean (EMK) exhibited highly productive vowel harmony, including harmony within word stems and harmonic alternations for many of its case particles, these rules have become less productive and applicable to fewer phonological contexts from Late Middle Korean (LMK) onward. Language-internal changes have been correctly identified as one motivator for these rule reductions, yet the role of the extensive Sino-Korean lexicon in sound changes in LMK has not been thoroughly explored. Recent work in usage-based functionalism indicates that speakers infer and ultimately construct the rules of their language by inferring patterns from their accumulated linguistic data and experiences. It is argued here that as the local schema of non-harmonic Sino-Korean loans grew in prominence, these loans began to interfere with native speakers’ inference of earlier harmonic rules, and said rules were relaxed; Sino-Korean loans were not merely exceptions to existing harmonic rules but catalysts of broader harmonic change.

Key words

Sino-Korean, Usage-Based Functionalism, Vowel Harmony, Cognitive Linguistics
1. Introduction
As with numerous languages of Central and Northeast Asia, historical and modern, Korean has long had a harmonic paradigm based on alternations between yin/dark and yang/light vowels. Where it differs from many Turkic and Mongolic languages, as well as Old Japanese (Hattori 1982), is that Korean harmonic alternations, at least from Middle Korean onward, do not exhibit a clear paradigm in the manner of [+/-back], [+/-open], or the like. (Old Korean is perhaps best analyzed as having had a [+/-RTR] [retracted tongue root] contrast; see Ko, Joseph, and Whitman 2014:143.) The harmonic paradigm observable in Table 1 below has been in place largely from Late Middle Korean (LMK) through Contemporary Korean (CK).

<table>
<thead>
<tr>
<th>Yang</th>
<th>ㅏ</th>
<th>ㅏ</th>
<th>ㅗ</th>
<th>ㅗ</th>
<th>ㅐ</th>
<th>ㅐ</th>
<th>ㅚ</th>
<th>ㅚ</th>
<th>ㅡ</th>
<th>ㅡ</th>
<th>ㅣ</th>
</tr>
</thead>
<tbody>
<tr>
<td>헤</td>
<td>a, ya</td>
<td>-o, yae</td>
<td>ae, yae</td>
<td>oe, wae</td>
<td>ɔ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yin</th>
<th>ㅓ</th>
<th>ㅓ</th>
<th>ㅜ</th>
<th>ㅜ</th>
<th>ㅔ</th>
<th>ㅔ</th>
<th>ㅟ</th>
<th>ㅟ</th>
<th>ㅡ</th>
<th>ㅡ</th>
<th>ㅣ</th>
</tr>
</thead>
<tbody>
<tr>
<td>헤</td>
<td>eo, yeo</td>
<td>u, yu</td>
<td>e, ye</td>
<td>wi, wa</td>
<td>we</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neutral</th>
<th>-eu</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Basic Harmonic Paradigm in Korean

One problem of interest to both diachronic linguistics and cognitive/psycholinguistics is that vowel harmony in Korean was more productive in Early Middle Korean (EMK) than in the current language (Kim 1978; Chang 1982; Park 1990; Sohn 2001; Lee and Ramsey 2011). For instance, native Korean lexemes as of EMK featured stem-internal harmony by default, whereas in CK this is only an active process in the context of sound-symbolic vowel harmony (Park 1990), in which the stem-internal vowels of stative verbs and onomatopoeia alternate to reflect different connotations or shades of meaning. Likewise, both EMK and CK feature harmonization of initial vowels in many verbal suffixes and case particles with the stems to which they are attached, yet EMK exhibited harmony for a larger selection of case particles. This can be seen with the topic-marking case particle -eun/-neun -은/는, which does not alternate harmonically in CK yet EMK exhibited the realizations -ɔn/-nɔn (Park 1990; Lee and Ramsey 2011). Another instance is the case particle -eu.lo -으로 ‘with,’ which previously alternated with -ɔ.lo. Several more examples of native EMK harmony follow are portrayed in Table 2 below.

1 Transcriptions here use the McCune-Reischauer romanization scheme, which align with IPA representations of vowels with the following exceptions: Grapheme <y> is /j/, <ae> was historically /ɛ/ but has merged with <e>/e/, <eo> is /ʌ/ or /ə/, <eu> is /ɪ/, <oe> is /we/ or /ə/, <wi> is /wi/ or /y/, and <ui> is /ɨj/ or /i/.
2 Denotes a historical vowel that has since split into the realizations <ㅏ>/a/ and <ㅡ>/i/.
3 <eu> is harmonically neutral in non-initial syllables within the context of sound-symbolic (word-internal) harmony in Contemporary Korean (Kim 1977). In other contexts, it functions as a yin vowel.
Table 2. Native EMK Harmony

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>Translation</th>
<th>Type of Harmony</th>
</tr>
</thead>
<tbody>
<tr>
<td>버들</td>
<td>peo.teul</td>
<td>Willow</td>
</tr>
<tr>
<td>구름</td>
<td>ku.rum</td>
<td>Cloud</td>
</tr>
<tr>
<td>사람</td>
<td>sa.ram [MK sa.ɾam]</td>
<td>Person</td>
</tr>
<tr>
<td>벌어</td>
<td>peol-eo</td>
<td>to earn</td>
</tr>
<tr>
<td>살아</td>
<td>sal-a</td>
<td>to live</td>
</tr>
<tr>
<td>먹어</td>
<td>meok-eo</td>
<td>to eat [informal]</td>
</tr>
</tbody>
</table>

Being that Korean has been written (at times) in its own alphabet since the 15th century, both vowel harmony and a substantial number of instances of disharmony have been documented throughout the LMK period onward. The same period of Korean history also saw an influx of spoken as well as written Chinese loanwords enter the language, and it is tempting to endeavor to explain Korean harmonic change mainly within the context of language contact. A principled analysis of diachronic harmonic change, however, should first examine language-internal factors that may have steered the course of Korean phonology.

1.1 Established factors in Middle Korean phonological change

The split of the historical mid-back yang vowel < ㅗ > /ɔ/ beginning in the LMK period occurred within the context of what Chang (1982) describes as a tumultuous sociolinguistic situation: the 14th and 15th centuries in the Korean peninsula saw a regime change with the founding of the Joseon Dynasty, “the weakening of the early feudal structure [in Korea],” and the invention of the Hangul alphabet (134). Chang argues at length that the sociopolitical changes in the Peninsula resulted in unusual social mobility for the era, enabling an intermingling of relatively conservative dialects of Korean with linguistically innovative ones. This situation likely allowed for the above-mentioned vowel split throughout the Korean peninsula.

The change under discussion here is that of the shift of yang vowel < ㅗ > /ɔ/ to yin < ㅡ > /ɨ/ in non-initial syllables, a change that was completed during the 16th century. 4 (< ㅗ > /ɔ/ also shifted to < ㅏ > /a/ in initial syllables beginning in the 18th century, but that does not constitute a harmonic change and is therefore less relevant to this analysis.) Ko (2012:185) identifies an /ɨ/-regressive [partial] assimilation during the 15th century in which nominals featuring /ɨ/ in the final syllable saw vowel-raising in the previous syllable, such as /cjə.pi/ ‘bird’ to /cje.pi/ and /kil.ək,i/ ‘wild goose’ to /kijl.ək,i/. Crucially, Lee and Ramsey (2011:173) find the form /kil.ək/ for ‘wild goose’ in Hunmin Cheongeum Haerye, a text from 1446, indicating that the final /ɨ/ in /kil.ək/i/ was added later as a pleonastic suffix. While Ko’s analysis focuses on assimilation within morphologically simple words, the case of /kil.ək/ to /kil.ək,i/ and then to /kil.ək,i/ demonstrates raising caused by the addition of a suffix with /ɨ/. Thus, the raising of non-initial

4 The phonetic distance between /ɔ/ and /ɨ/ might not be as great as it seems. The transcription /ʌ/ is favored by Ko (2012) and others for the vowel < ㅗ > /ɔ/, and /ɨ/ is sometimes transcribed as /u/ in the literature.
vowels in such environments needs not be limited to nominals as Ko (2012) states; indeed, Lee
and Ramsey (2011:145, 177) identify a causative/passive suffix, productive during the 16th
century, with the allomorphs /-hi/, /-ki/, /-Gi/\(^5\), and /-i/. These findings present plausible
scenarios for raising of non-initial vowels—perhaps /ɔ/ to /ɨ/, for instance—in non-initial
syllables.

This change proves disruptive to both stem-internal harmony and stem-case particle/stem-
suffix harmony. Examples of the former include /ta.Łɔ/ ‘different’ to /ta.Łɨ/ (Sohn 2001), /a.Łɔˈl/
‘son’ to /a.Łi/ (Lee and Ramsey 2011), /mɔ.ɔm/ ‘mind’ to /ma.i.mi/ (Kim-Renaud 2008);
examples of the latter include the previously-mentioned topic marker, the accusative marker -
eul/leul -을/를 with its previous realizations of -ɔ/lɔl, and the previously-mentioned -eu.lo -으로
‘with.’ In all of these instances, a word stem or a word and its attached suffix/particle no longer
agree harmonically. The widespread nature of this change has surely contributed to the reduced
productivity of harmony in CK.

1.2 Understanding the role of Sino-Korean loans in harmonic change
All of this being said, arguments such as those posed by Chang (1982), Kim-Renaud (2008), and
Lee and Ramsey (2011) give credence to the notion that the Sino-Korean lexicon, having
expanded throughout the Middle Korean period, was sizable enough in LMK to have had an
influence on the broader phonology of the language. After all, due to the shared history between
the two language groups and China’s long-term political and cultural hegemony in East Asia, the
body of Korean vocabulary composed of Sino-Korean loans is both extensive—likely 60 to 70%
of the entire lexicon (Sohn 2001; Jung and Cho 2006)—and diverse in terms of semantic
categories. Further, it has been established that native speakers of a language, while broadly
aware that their native tongue has borrowed from other languages, are not always aware of which
words in their lexicon are borrowings. It is well understood, for instance, that the modern English
lexicon hosts an abundance of recognizable French loans such as restaurant, chaise, chef and
chauffeur, but the fact that words such as chair and chef are older loans from French is likely lost
on numerous native English speakers. The Sino-Korean lexicon can be characterized similarly.
While some borrowing has taken place throughout the history of relations between Korean and
Chinese speakers, according to Lee and Ramsey (2011), the earliest borrowings “do not look like
Sino-Korean and are not thought of as Chinese; they are considered native and never written
with Chinese characters” (76), and “only scholars know that they are of Chinese origin” (69). As
for those Sino-Korean loans that are widely acknowledged as such, Eom (1999) emphasizes that
the phonology of those lexical items contains “traces of Old Chinese, Middle Chinese, Old
Mandarin, and Modern Mandarin,” though it is largely derived from Middle Chinese (27).
Adding to this, Lee and Ramsey point out that Korean received an influx of Chinese loans which
were “apparently borrowed directly from spoken Chinese” during the 16th century and reflect
northern Chinese pronunciations (2011:237-8). It is worth considering whether this extensive
borrowing has gone beyond contributing exceptions to Korean harmonic rules, coloring Korean
vowel harmony itself.

The theoretical school of generative phonology has seen numerous attempts to systematically
characterize vowel harmony, particularly for languages of Central Eurasia ranging from
Hungarian to Mongolian and Classical Manchu. Unfortunately, though synchronic, generative

\(^5\) /G/ represents a Middle Korean consonant, generally velar in realization, which later underwent lenition and
deletion (Lee & Ramsey 2011:10).
Section 2 below describes the relevant linguistic features of the Sino-Korean lexicon and likely avenues of influence on Korean phonology. Section 3 summarizes findings from UBF and demonstrates how these findings further support the notion that the Sino-Korean loans were not merely exceptions to the existing harmonic ruleset but in fact motivated changes to this ruleset at the psycholinguistic level.

2. Linguistic features of Sino-Korean loans: Influences on Middle Korean harmony

The Korean language has retained conservative pronunciations of its Sino-Korean loans, as can be seen instances such as the Chinese character 想 and its regular Korean pronunciation 상 /saŋ/; the same goes for the character 念 and the corresponding pronunciation 넣 /njəm/. Further, Lee and Ramsey (2011) cite examples of Chinese morphological suffixes that were attached to native Korean words irrespective of harmony. One example is the disharmonic /twit.kan/ ‘toilet,’ -간 /-kan/ being derived from Sino-Korean 間 ‘room.’ Another is /mə.zəms.kjəŋ/ ‘state of mind,’ in which -경 /-kjəŋ/ derives from Chinese 景 ‘view, scenery.’ These internal correspondences and the violations of the normative LMK harmonic rules are, in fact, consistent with findings from prior studies of the phonology of loaning as demonstrated in Korean. Itō et al. (2006), Kang (2003), and Kim (2008) all find that loanword adaptations constitute attempts to faithfully

6 Parts of Sections 3 and 4 are adapted from earlier work in de Roulet (2018).
represent the original pronunciation of the words in question, at least as far as first-language (L1) perception and articulatory phonetics allow. Their claims are supported by earlier research (Silverman 1992) on English loans in Cantonese indicating that unfamiliar linguistic inputs (i.e. loans from a language in which listeners are not proficient) are received as acoustic inputs, not strictly linguistic ones, which listeners process and segment within the bounds of their L1 phonology. The adherence of Sino-Korean loans to their earlier word forms remains robust, even in defiance of Middle Korean harmonic rules. A list of Sino-Korean words provided by Jung and Cho (2006) presents 104 instances (out of a total of 261, or 39.8%) of Sino-Korean loans that stand in violation of stem-internal or stem-suffix/case particle harmony. (In the remaining cases, the original Chinese pronunciations happen to be compatible with Korean harmonic rules.) Several examples follow may be seen in Table 3.

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>Translation (from Korean)</th>
<th>Harmonic Rule Violated</th>
</tr>
</thead>
<tbody>
<tr>
<td>악명 (惡名) ak.myeong</td>
<td>notorious; notoriety</td>
<td>stem-internal</td>
</tr>
<tr>
<td>상념 (想念) sang.nyeom</td>
<td>Thought</td>
<td>stem-internal</td>
</tr>
<tr>
<td>사유 (思惟) sa.yu</td>
<td>reason (i.e. justification)</td>
<td>stem-internal</td>
</tr>
<tr>
<td>도서관 (圖書館) to.seo.kwan</td>
<td>Library</td>
<td>stem and derivational suffix</td>
</tr>
<tr>
<td>뒷간 ( - 間) twi.kan⁷</td>
<td>Toilet</td>
<td>stem and derivational suffix</td>
</tr>
<tr>
<td>백두산 (白頭山) paek.tu.san</td>
<td>Mount Paekdu</td>
<td>stem-internal</td>
</tr>
</tbody>
</table>

The instances in which harmonic rules are violated by the joining of a derivational suffix to a harmonically mismatched stem are particularly noteworthy (to.seo.kwan ‘library’ and twi.kan ‘toilet’ above). That native speakers of a language can reanalyze loanwords as being original to their language has already been established. However, it is also worth considering whether lexemes with nonharmonic (Chinese) derivational suffixes attached might be reanalyzed as being morphologically simple, then presenting further exceptions to the apparent rule of stem-internal harmony as seen in EMK. This issue is discussed further in Section 4.2.

What remains is the question of whether these nonharmonic Sino-Korean loans would have been taken as isolated (lexicalized) exceptions to Korean harmonic rules or would have had an influence on these harmonic rules more broadly. As diachronic linguistic changes are necessarily transferred from one speaker to another and stored in the memories of learners, the field of cognitive linguistics is well-equipped to address this issue.

3. Usage-Based Functionalism: Applications for analysis of Middle Korean harmony
3.1 Fundamental claims of usage-based functionalism
Usage-based functionalism (UBF), a theoretical school of cognitive linguistics with precursors in the functional-typological approach of Greenberg (Bybee 2013), holds that linguistic knowledge (competence) and experience (performance) are intertwined and inseparable, and that the

---

⁷ 뒷 twi.kan is a native Korean word meaning ‘rear’; 간 (間) here derives from Sino-Korean -kan/ ‘room’ (Lee and Ramsey 2011:175).
structure of one’s language is inferred from the data one receives. Bybee (2006), an author of numerous works on usage-based phonology and the influence of frequency effects on linguistic change, also asserts that diachronic language change “plays an indispensable role in the formulation of explanations for linguistic patterns.” This makes the work of Bybee and others in UBF naturally well-suited to explaining the reduction of harmonic productivity—a well-attested historical phenomenon—through the history of Korean. Of particular relevance to this analysis are UBF-based findings on phonology, the impact of frequency effects on L1 acquisition, and network morphology.

3.2 Usage-based phonology and the inference of harmonic rules

Given Kang (2003), Kim (2008), and Silverman’s (1992) work on the perceptual and phonetic constraints governing loanword phonology, there remains a question of how it is the case that Sino-Korean loans have retained relatively conservative pronunciations rather than changing to fit MK harmonic rules. Bybee’s (1999) distinction between global rules and local schemas is useful here: one key finding from UBF is that “local schemas take precedence over general ones” (1999:211)—a logical claim, as rule-exceptions within a language would not remain stable otherwise. High frequency of use also reinforces a given form and aids in retaining it, even in the those which prove to be exceptions to the broader rules of a language (Bybee 2006, 2013; Kapatsinski 2014). Some of these schemas in Middle Korean would have consisted of the language’s Sino-Korean lexemes, which, as stated above, exhibit consistent correspondences with the original Chinese. The nonharmonic forms of these loans could have then been further reinforced by association with each other through their similar phonological shapes; loanwords such as 도서관 (圖書館) to.seo.kwan would retain their forms via a combination of association with other words with similar forms and high frequency of use. Relatedly, words featuring Sino-Korean derivational suffixes, such as 뒷간 ( - 間) twi.kan, may be grouped together according to their shared morphology (see Section 4.3) or reanalyzed as monomorphemic lexemes, in which case they, too, appear to be exceptions to the rule of stem-internal harmony.

UBF also provides a model for the manner by which the accumulation and retention of these rule-exceptions can interfere with the inference of the language’s global or broadly productive rules. Bybee (2006) asserts that grammar is “the cognitive organization of one’s experiences with language” and that “the most general phonological analysis is not the one speakers necessarily use” (Bybee 1999; see also Tomasello 2009). Rather than having preset rules through which they analyze and process language, learners—L1 or otherwise—first receive linguistic data and learn forms at the local level, and over time they observe apparent patterns from this extensive input. This process for vowel harmony rule acquisition in EMK is modeled below:
In EMK, with its comparatively smaller number of Sino-Korean loans, a rule of stem-internal harmony is readily inferred by sorting the available linguistic data; nonharmonic lexemes are learned as exceptions to rules that are otherwise applicable to much of the language and can be relied upon for forming novel utterances that are deemed grammatical by other speakers. As the usage of (frequently non-harmonic) Sino-Korean loans increased during the 15th and 16th centuries, however, the pattern of stem-internal harmony would be less obvious in perception and thus less productive in usage. Further, given that speakers do not always draw distinctions between loanwords and native lexemes, historical cause and effect behind the harmony of some lexemes and the disharmony of others may not be apparent. The mixed data received by speakers may, at the level of an entire language, result in a relaxing of rules over time when viewed diachronically. (Cases such as 나모 na.mo, spelled and pronounced 나무 na.mu in CK, may be instructive.)

3.3 Morphological networks and harmonic suffixes/case particles
UBF diverges from generative analysis in its treatment of morphology as well as in its coupling of linguistic knowledge and experience. Rather than assuming that morphemes are the basic units of analysis in morphology, UBF asserts instead that language can be learned in larger chunks, those being words or in some instances entire phrases, with Tomasello (2009:325) stating that “the most fundamental units” of language are “whole utterances and constructions—not isolated words or morphemes.” In this conception, individual high-frequency words which fit only a local-level rule or none at all can also be “stored together with bound morphemes as prefabricated units” (Diessel 2017). This notion that a word can be stored and recalled as a whole, as opposed to being assembled from its morphemes when used in speech, helps to explain the reanalysis of morphologically complex words as being morphologically simple. One such instance is the Sino-Korean 도서관 to.seo.kwan ‘library’, originating from the morphologically complex 圖書館 (圖書 itself being a morphologically complex pair of characters meaning ‘books’ and 館 meaning ‘building’ or ‘center’) but likely morphologically simple to most speakers of Korean.

Bybee claims, further, that words are stored in morphological networks, systems of word-association in which the words’ commonalities in features assist with their recall. Diessel (2017) illustrates the morphological network model below:
This model helps to explain the acquisition and retention of regular harmonic forms such as many Korean verb endings, yet it also sets a precedent for the partial recall of conjugations of irregular verbs such as 이다 i.ta ‘to be.’ Not only is 이다 a sufficiently high-frequency verb that it can be easily stored in the mental lexicon, but recollection of its conjugations (i.e., 이에요 /예요 i.ae.yo / yae.yo in a polite register and 입니다 ip.ni.ta for more formal contexts) is assisted by association with regular verbs sharing the -요 and -ㅂ니다 endings.

A morphological network could also enable the cognitive association of morphological blends featuring Sino-Korean derivational suffixes, such as 뒷간 twi.kan ‘toilet’ with other words ending in the Sino-Korean 간 (derived from Chinese 間). What Bybee’s model contributes to a UBF-based analysis of Korean harmony is a means by which non-harmonic local schemas are able to be recalled with relative ease, enabling these local schemas to persist despite phonological regularity of numerous other Korean lexemes. Given this possibility of retaining
local-level exceptions to apparent rules, it is also possible for a sufficient number of local-level exceptions to accumulate in a speaker’s mental lexicon that far-reaching patterns are not inferred as easily by the speaker.

4. Discussion and conclusion

The advantage of analysis via UBF for characterizing harmonic change is ultimately that the models provided by UBF answer questions which generative/synchronic analysis and even diachronic analysis alone do not. Whereas diachronic analysis is popular among Altaicists and other linguists interested in vowel harmony, and indeed it often effectively models changes in a language’s features over its history, diachronic analysis alone does not explain the means by which language change occurs at the psychological level. Meanwhile, the work of linguists such as Bybee in phonology and Tomasello in language acquisition explains how harmonic rules can be constructed from the linguistic data available to speakers and how said rules can be reanalyzed or fail to be passed on to future generations.

The analysis of the circumstances of Middle Korean is not a mere theoretical exercise; Contemporary Korean faces a tumultuous sociolinguistic situation today much as it did five centuries ago. For one, it would be worth analyzing the available documentary evidence for Early Modern Korean to see how the recent Peninsula’s history of interaction with and occupation by Japan further shaped or eroded harmony in Korean. Certainly, Contemporary Korean hosts a considerable number of loans from Japanese, not to mention secondhand English loans via Japanese (Itô et al. 2006, Kim 2008). In the aftermath of Japanese occupation, of course, social mobility and human migration in the Peninsula have grown with astounding speed, with South Korean economic growth decelerating only over the past several years. Further, to say that proficiency in English has become socially desirable in South Korea due to the country’s entry into the global market is perhaps too great an understatement. There is even a question of whether China, South Korea’s largest trade partner, will once again become an important source of cross-linguistic influence given its growing economy, its investment in the sciences, and the rise of its universities in international rankings. To analyze Korean harmonic change through the lens of UBF, then, is to potentially offer timely insights into the future of the Korean language at large.

References


