

Description and Explanation in Inflectional Morphophonology: The Case of Korean Noun Stems

Brent de Chene

This study is part of a larger project that has two aims. The first is to use ongoing change in a morphophonological system to infer the descriptively adequate analysis of that system—that is, the analysis that speakers have adopted in favor of other observationally adequate analyses. The second is to use comparison between the descriptively adequate analysis and other observationally adequate analyses to infer the principles by which the descriptively adequate analysis has been chosen and in this way approach an explanatorily adequate theory of inflectional morphophonology. The paper will be divided into two sections that deal, respectively, with the problems of description and explanation in the domain of Korean nominal inflection. Throughout, a point of reference will be the analysis of Japanese verbal inflection published as de Chene 2009.

1. Description

The apparent reanalysis, in Korean, of noun stems in $t^h c c^h$ as s -stems, ongoing for t -stems and c -stems in particular since at least the 18th century (Kang 2003: 129–130), is a change that displays both interesting similarities and interesting differences with respect to the Japanese change analyzed in de Chene 2009. In particular, the Korean change is similar to the Japanese one at the descriptive level in that, because one of a set of alternations has been chosen as regular, the elimination of exceptional allomorphs results not in leveling but in the exchange of an irregular alternation for the regular one. It is dissimilar at the explanatory level, I will claim, in that the conjunction of Hypotheses 1 and 2, repeated below, is insufficient to account for it.¹⁾

Hypothesis 1

Base forms are surface alternants drawn from a morphologically or phonologically defined environment that is fixed over the entire range of the alternation and over an entire lexical category.

Hypothesis 2

Base forms of both stems and affixes are chosen, consistent with Hypothesis 1, so as to allow prediction of the maximal number of inflected forms from the paradigm of the maximal number of stems.

The relevant Korean alternations are the result of the fact that while the eight coronal obstruents $t^h t' c c^h c'$ $s s'$ contrast in syllable onsets, the only coronal obstruent permitted in syllable codas is t , realized as $[t^ʔ]$. A

stem-final consonant in Korean will be a coda when followed by pause, by a consonant, or by a vowel across a boundary that blocks resyllabification, in particular a word or compound boundary. It will be an onset when followed by a vowel across a boundary that does not block resyllabification, in particular an affix or clitic boundary, where clitics include case particles and the copula (see Martin 1954:20, Sohn 1999:165). Of the eight coronal obstruents, t' c' do not occur stem-finally. We thus expect the following six alternations depending on whether a stem-final coronal obstruent finds itself in a coda (left-hand alternant) or, as the result of resyllabification, in an onset (right-hand alternant; allophonic voicing of plain oral stops and affricates in voiced environments is disregarded).²⁾

- (1) a. $t \sim .t$ c. $t \sim .c$ e. $t \sim .s$
 b. $t \sim .t^h$ d. $t \sim .c^h$ f. $t \sim .s'$

In verb stems, all six of the alternations of (1) occur and are stable. In noun stems, the first five occur, but only the fifth, (1e), is stable: the other four all show a tendency to be replaced by the fifth.

In the treatment of Japanese verbal suffix alternations in de Chene 2009, predictions concerning stability and instability of suffix alternants are computed for three possible analyses, reasoning from synchronic structure to potential change according to the first half of Kiparsky's (1978 [1982:217]) dictum concerning the relation between synchronic structure and diachronic change, namely that structure constrains change; those predictions are then compared with the facts of change in progress. With regard to Korean, we will reason in the opposite direction according to the second half of Kiparsky's dictum, namely that change can diagnose structure. The form of morphological change predicted by our model (de Chene 2009) of how basic, derived, and irregular morpheme alternants are accessed and generated is the replacement of irregular nonbasic forms by regular derivatives, reducing to replacement of marked by unmarked alternants in the absence of a rule. In a situation where alternant type X is stable, but alternant type Z is subject to replacement by alternant type Y, then, it follows that type X is basic, type Z is irregular, and type Y is generated by rule from type X.

For example, given that in Korean, pis is replacing pic as the resyllabification environment (below, "non-neutralized") alternant of pit "debt", which is stable as the "neutralized" alternant, we infer that pic is irregular, pis is a regular derivative, and $/pit/$ is the base form. This in turn entails that there is a rule taking t to s when t is both final in a noun stem and syllable-initial (i.e. resyllabified), as in (2).

- (2) $t \rightarrow s / _ _ N]$

Postulating rule (2) along with representations in which alternants ending in t , t^h , c , and c^h are irregular (i.e. marked and lexically listed) provides an account for why, of the five alternations (1a)–(1e), only the last is stable in noun paradigms and why, in the history of the language, nouns that have lost alternants ending in those four consonants have acquired alternants ending in s as a result. It also provides an account for the fact that when a noun is borrowed in a form ending in t , that t invariably alternates with s before a case particle or the copula (Kang 2003:130, Kenstowicz and Sohn 2001: 265–266).

Just as t is the only coronal obstruent allowed in syllable codas in Korean, p and k are the only labial and velar obstruents allowed in that environment; more generally a coda consonant must be one of the set $p t k m n$

ŋ l. Since historically there are noun stems ending in $p^h k^h k'$ and in the consonant clusters *ps ks ls lp lk*, we observe the following alternations, parallel to those of (1), depending again on whether a stem-final consonant (cluster) finds itself in a coda (left-hand alternant) or (at least partially) in an onset (right-hand alternant).

- (3) a. $p \sim .p^h$ c. $k \sim .k'$ e. $k \sim k.s$ g. $l \sim l.p$
 b. $k \sim .k^h$ d. $p \sim p.s$ f. $l \sim l.s$ h. $k \sim l.k$

Below, we will refer to noun stems ending in clusters or in single consonants other than *p t k m n ŋ l* as having “noncanonical” finals.

If neutralized alternants are basic for Korean nouns whose stems end historically in coronal obstruents, as we have concluded above, we expect the same to be true for nouns in general, in accordance with the provision of Hypothesis 1 that base forms are chosen from an environment that is fixed over the entire range of the alternation and over an entire lexical category. For nouns ending in clusters, this expectation is confirmed by a long-standing tendency to generalize the neutralized alternant to resyllabification-triggering environments. Thus, for example /kaps-il/ “price (acc.)” and /talk-il/ “chicken (acc.)” are typically realized as /kap-il/ and /tak-il/, respectively (Martin 1992: 107–108). For nouns ending in $p^h k^h k'$, the tendency to generalize the neutralized form is weaker, but still apparent: Martin (1992:108) cites /puək/ for /puək^h/ “kitchen” as one well-known example.

In sum, the descriptively adequate analysis of Korean noun stem alternations appears to be that (a) neutralized forms are basic and (b) stem-final *t* is taken to *s* syllable-initially by rule (2). The resulting classification of alternants for the three stems “debt”, “kitchen”, and “price” is shown in Table 1.

Table 1 Irregular/Basic/Derived Alternants of Representative Noun Stems

Irregular	Basic	Derived	
pic	pit	pis	“debt”
puək ^h	puək		“kitchen”
kaps	kap		“price”

For “debt”, loss of the irregular alternant results in the appearance of the regular derivative; for “kitchen” and “price”, loss of the irregular alternant results in leveling in favor of the base.

For verbs, in contrast to nouns, there is no evidence that neutralized stem alternants are basic. I will assume that non-neutralized (resyllabification environment) alternants of verb stems are basic, and correspondingly that for verbs, alternations like those in (1) and (3) are the result of rules of cluster simplification and coda neutralization (see e.g. Ahn 1985: 166–173). The coexistence in the grammar of a rule of coda neutralization and the inverted *t*-to-*s* rule (2) qualifies this as a case of “synchronic rule inversion” (McCarthy 1991).

2. Explanation

The two elements of the analysis of noun stems just proposed, and thus the two explicanda of that analysis, are the choice of neutralized over non-neutralized alternants as basic and the choice of the *t*-to-*s* rule (2) from among five possibilities (i.e. those corresponding to (1a)–(1e)) as the rule governing the alternations of coronal obstruents. Given that *s* is the most common stem-final coronal obstruent (at 56% of the total, more than twice as common as second place *cʰ* in Kim and Kang 2000 (Albright 2008:171)), it seems clear that the latter choice is governed by predictability considerations in accordance with Hypothesis 2 (as suggested, for example, by Kang (2003: 131)). Since Hypothesis 1 would be satisfied by taking either neutralized or non-neutralized alternants as basic, an obvious possibility is that the choice of neutralized alternants as basic is determined by Hypothesis 2 as well.

Essentially this possibility has been pursued by Albright (2008), who claims that three factors conspire to make neutralized alternants the more reliable choice. Crucially, however, Albright assumes that speakers calculate basic forms not of morphemes, but of paradigms, and that the basic forms themselves are thus not morpheme alternants, but inflected forms. In this connection it should first be noted that there is in fact nothing that can confidently be called a nominal paradigm in Korean: the particles (Korean *cosa*) that might be thought to define a notion of paradigm are (1) categorially promiscuous, attaching to a variety of phrase types up to and including TP/CP (see Sohn 1999: 213) and (2) ill-defined as a class, shading off from structural case markers through markers of discourse function to exponents of a variety of spatial, temporal, and logical (“each”, “only”) meanings. Structural case markers, as is often noted, are also unaffixlike in being deletable. Abstracting away from the inapplicability of the concept “paradigm”, however, let us examine the reasoning behind Albright’s claim that the neutralized alternant—in his terms, the unsuffixed or unmarked member of the nominal paradigm—are taken as basic on grounds of predictability.

Albright first notes (2008: 171–172) that there are so few noun stems with noncanonical finals that, across the entire lexicon, taking the unsuffixed form as basic will result in only a minor predictability deficit relative to taking as basic the nominative in *-i/-ka*.³ (Taking the nominative as basic will also involve a small predictability deficit because of a rule merging dental with palatal stops before *i*). He next claims (2008: 172–173) that taking the accusative in *-il/-il* as basic involves a predictability deficit greater than that associated with the unsuffixed form because a sequence *XVil* is ambiguous between an analysis /XV-il/ and an analysis /XV-i-l/. Finally, he claims (2008: 173–175) that case markers are suppressed so often, at least in child-directed speech, as to make the nominative a less reliable base than the unsuffixed form on the grounds of (token) frequency. His conclusion is that the unsuffixed form is the most reliable base.

In summarizing the lessons of his investigation of Japanese verbal morphology, de Chene (2009) notes that the interparadigmatic nature of the alternations involved makes it impossible to regard as a general theory of inflectional morphophonology any analytic model according to which inflected forms are derived from other inflected forms within a paradigm. He concludes that, as long assumed by structural and generative linguists, the morpheme rather than the inflected word is the basic unit of morphophonological analysis. Applying this lesson to Korean noun stem alternations, we see that the question posed by those alternations is not that of which inflected form (i.e. stem-clitic combination) is basic with respect to other inflected forms within the paradigm, but that of which stem allomorph is basic with respect to other stem allomorphs. One consequence of taking morphemes rather than words to be the basic units of morphophonology is that the segmentation of

word forms into morphs can be assumed as a precondition for morphophonological analysis, just as segmentation of phrases into words is typically assumed without comment in a word-based system.

In this context, consider the ambiguity of a sequence *XVil*. This is precisely a problem of segmentation of word forms into morphs, and it can be resolved by the comparison of that sequence with any other sequence consisting of the same stem plus a different suffix or clitic. For example, if the result of adding the topic clitic *-in/-nin* to the stem in question is *XVin*, the analysis of the original sequence is /XV- il/, whereas if the result of adding the topic clitic is *XVnin*, the analysis of the original sequence is /XV-lil/. The ambiguity of *XVil* is parallel to the ambiguity, in Japanese, of a verbal indicative form like *iru* between /i-ru/ and /ir-u/; other inflected forms provide disambiguation, and the distinction is stable, indicating that learners find it transparent. I conclude that there is no morphophonological predictability deficit associated with the ambiguity of segmentation posed by *XVil*: segmentation into morphs is unproblematic for speakers and is logically prior to morphophonological analysis in any case.

Next consider the question of whether the deletability of the nominative and accusative case markers is likely to pose a problem of rule reliability, as Albright claims. Albright is of course concerned about the reliability of rules deriving the unsuffixed form from one or another inflected form; this is an issue that has no direct counterpart in a morpheme-based framework, where the coda neutralization rule that would derive neutralized from non-neutralized alternants if the latter were basic (as we have concluded they are for verbs) is automatic and surface true. Even in a word-based framework, however, it is unclear that the unreliability of rules based on the nominative (or accusative) can explain why neutralized alternants are basic. This is because there are five more clitics with high token frequency that create resyllabification environments for noun stem finals and for which stem-clitic combinations thus constitute further candidates for base forms.⁴⁾ The entire set of seven clitics that have this property is given in Table 2; frequencies, from Kim and Kang 2000, include minor allomorphs and sequences that begin with the morpheme in question and are rounded to the nearest thousand (excluding symbols, the corpus in question contains 3,036,891 morphemes, so that the 90,000 of the Nominative, for example, represent about 3% of the total).

Table 2 Clitics Triggering Resyllabification of Noun Stem Finals

Category	Form	Frequency
Copula	-i-	65,000
Topic	-in/-nin	86,000
Nominative	-i/-ka	90,000
Accusative	-il/-lil	114,000
Genitive	-e (- <i>ty</i>)	84,000
Dative	-e	87,000
Instrumental	-ilo/-lo	44,000

Concluding on the basis of one or more of the frequently deleted structural case markers that there is no reliable base other than the neutralized alternant would thus seem premature.

We have seen that the issue of rule reliability that concerns Albright does not have an exact counterpart in a morpheme-based framework. There is, however, a question we can ask in such a framework about the information available to speakers that is roughly parallel to that of rule reliability. That question is whether children encounter enough instances of the prevocalic alternant of a noun stem to set that alternant up as basic assuming that other conditions permit that choice. While there is to my knowledge no established standard for deciding this question, suggestive evidence is provided in the Korean case by a comparison of nominal and verbal inflection.

Many of the most common verbal suffixes of Korean are invariably consonant-initial, with the result that forms in which those suffixes attach directly to verb stems provide no information about the identity of noncanonical stem-final consonants. This is particularly striking in main clause forms of the “plain” verbal paradigm, which, crucially, is the speech level used by adults to children and by children among themselves up through high school (Lee and Ramsey 2000: 253–254): of the six forms of that paradigm lacking medial suffixes, only the imperative (and, for stative verbs, the interrogative indicative) is formed with a suffix that takes a vowel-initial form after a consonant-final stem (see Sohn 1999: 236–237). Further, the most common verbal suffixes that do begin with vowels, either in general or after a consonant-final stem, are no more frequent than the clitics of Table 2; the figures from Kim and Kang 2000 for past tense *-a/əsʰ-* and infinitive *-a/ə* are 66,000 and 35,000, respectively. It thus seems possible that children may have less evidence about the noncanonical finals of verb stems than about those of noun stems. Yet they clearly have no trouble computing stem-final consonants for verbs: as we have already observed, verb stems with noncanonical finals are stable, displaying no tendency to undergo restructuring. The indicated conclusion would seem to be that speakers’ failure to take the non-neutralized alternant of noun stems as basic cannot be attributed to any paucity of evidence concerning those finals.

We have concluded that the two considerations other than strict predictability over the set of inflected forms that are claimed by Albright to enter into the calculation of the base form for Korean nouns, the indeterminate segmentation of *XV/i/* and the loss of information due to deletion of structural case markers, are without effect, at least in the morpheme-based framework necessary for a general account of inflectional morphophonology. We are thus left with the original criterion of predictability. Let us look more closely at how that criterion applies to the present case.

To this point, we have been speaking as if Korean noun stems that end in a consonant have two alternants depending on whether or not the stem-final consonant is resyllabified as the onset of the following syllable. We have also noted, however, that dental stops become palatal before *i*. This means that for the approximately 0.3% of noun stems that end in *t tʰ* (see Albright 2008:171), there are two distinct resyllabification environment alternants, one occurring before *i*, the other before all other vowels. Let us call these two alternants the Pre-*i* and Pre-V forms, respectively, and extend this terminology to all noun stems.

Neutralization of the contrast between dental and palatal stops before *i* also means that for the approximately 0.7% of noun stems that end in *t tʰ c cʰ*, the Pre-*i* form will not be diagnostic of the Pre-V form. Let us say, then, that the choice of the Pre-*i* form as base has a predictability deficit of 0.7% and, correspondingly, a predictability index of 99.3%.⁵⁾ In contrast to the situation with the Pre-*i* form, the choice of

the Pre-V form as basic will not be associated with any predictability deficit, because there is no neutralization of consonants before vowels other than *i*. The Pre-V form will thus have a predictability index of 100%.

What will be the predictability deficit associated with the choice of the neutralized alternant as basic? To begin with, because a neutralized form ending in [tʷ] can correspond to a Pre-V form ending in any coronal obstruent (in practice, any of [tʰ cʰ s]), the neutralized form predictability deficit will be at least 1.5%, since this is the rough proportion of noun stems that end (originally) in coronal obstruents (666 of 43,933 in Kim and Kang 2000; see Albright 2008:171). In the same way, however, a neutralized form ending in [p] can correspond to a Pre-V form ending in any of [p pʰ ps], and a neutralized form ending in [k] can correspond to a Pre-V form ending in any of [k kʰ k' ks lk]. While the number of noncanonical stems is extremely small for final [p] and [k], the neutralized form will strictly speaking fail to be diagnostic of the Pre-V form in these cases as well—in sum, for any stem that ends in an obstruent.⁶⁾ Accordingly, another estimate for the predictability deficit associated with taking the neutralized form as basic is 18.5%, since this is the rough proportion of obstruent-final stems (8108 of 43,933 in Kim and Kang 2000).

Even on a conservative estimate, then, taking the neutralized form as basic will involve a predictability deficit twice as big as that associated with the choice of the Pre-*i* form as basic (1.5% versus .7%) and infinitely greater than that associated with the choice of the Pre-V form. It thus seems clear that, at least in a morpheme-based framework, it will not be possible to explain speakers' choice of the neutralized form as basic in terms of a higher degree of predictability associated with that choice. On the contrary, it will be necessary to appeal to a distinct criterion, one that is capable of overriding the predictability deficit we have seen to be associated with the neutralized form. Crucially, that criterion must enable us to distinguish between nouns and verbs, since, as we have seen, speakers show no tendency to reanalyze verb stems with noncanonical finals.

In fact, there is a clear criterion with precisely these properties, namely the occurrence of the neutralized alternant of noun stems before word boundary and prepausally, giving it the status of an isolation form. Verb stems, in contrast, never occur word-finally or, a fortiori, prepausally, and thus lack anything that could be called an isolation form.⁷⁾ While it is logically possible that the apparent preference for isolation forms as bases is derivable from more general principles, I will assume here that the preference in question is irreducible and must be written into our hypotheses about how base forms are chosen. In particular, I will assume that Hypothesis 1, which mandates that base forms are surface alternants drawn from a fixed morphologically or phonologically defined environment, must be supplemented by the condition that base forms coincide with isolation forms when those exist. Since we are assuming that the choice of base forms may not vary within a lexical category, this condition will not be counterexemplified by cases in which an isolation form exists for a proper subset of a lexical category but fails to be chosen as a base. Latin nouns, for example, do not in general have isolation forms, so the condition in question will be inapplicable to them. There will thus be no barrier, for an alternation like *honos* ~ *honor*-“honor”, to taking the isolation form *honos* not as basic, but as irregular; that it was so interpreted is shown by its elimination in the historical period (see Albright 2002, chapter 4).

Above we have seen that, as suggested at the outset, the reanalysis in progress displayed by Korean noun inflection is parallel to that displayed by Japanese verb inflection in that one of a set of alternations has been chosen as regular, so that the elimination of exceptional allomorphs results, at least in some cases, in the exchange of an irregular alternation for the regular one. At the explanatory level, however, we have seen that the conjunction of hypotheses we used to account for the Japanese reanalysis was insufficient to account for

- typically elided /V_C and desyllabified /V_V (Sohn 1999: 281), is preserved after a consonant.
- 5) These are the percentages of stems for which the entire “paradigm” —that is, the form the stem takes in all contexts— is or is not predictable on the basis of the pre-*i* form. Because of the indeterminacy of the notion “paradigm” in the case at hand, we are abstracting away from partial predictability within a “paradigm” —for example, from the fact that, in contrast to the pre-*V* form, the isolation form will always be predictable from the pre-*i* form.
 - 6) I will ignore the additional indeterminacy arising from the fact that [l] at the end of the isolation form can represent any of original /l ls lp/.
 - 7) While it is sometimes asserted that verb stems “require an inflection” —i.e., never occur unsuffixed (Kenstowicz and Sohn 2001: 255), a bare verb stem, like a bare noun stem, can appear as the first member of a compound; when this occurs, the compound boundary blocks resyllabification, as expected (puth- “stick” + an- “embrace” > putan- “hug”).
 - 8) This is arguably the account implicit in comments to the effect that speakers treat noun stems in /t th c ch/ “as if” they ended in /s/ (Martin 1954: 21 (fn. 12, 14), and it is apparently the account that would be endorsed by Hale and Reiss (2008), given their position that “Both paradigm leveling and proportional analogy can be attributed to the operation of a single mechanism—Lexical Restructuring.” (2008: 243) (Hale and Reiss (2008: 244–245) discuss Kenstowicz’s (1996) treatment of /kaps/ “price” but do not deal with stems ending in coronal obstruents.)

References

- Ahn, Sang-Cheol. 1985. *The Interplay of Phonology and Morphology in Korean*. Dissertation, University of Illinois. Published by Hanshin Publishing Co., Seoul.
- Albright, Adam. 2002. *The Identification of Bases in Morphological Paradigms*. Dissertation, UCLA.
- Albright, Adam. 2008. Explaining Universal Tendencies and Language Particulars in Analogical Change. In *Linguistic Universals and Language Change*, ed. Jeff Good, 144–181. Oxford: Oxford University Press.
- de Chene, Brent. 2009. Description and Explanation in Inflectional Morphophonology: The Case of the Japanese Verb. 早稲田大学教育学部学術研究 (英語・英文学編) 第56号.
- Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.
- Kang, Yoonjung. 2003. Sound Changes Affecting Noun-Final Coronal Obstruents. In *Japanese/ Korean Linguistics* 12, ed. William McClure, 128–139. Stanford: CSLI.
- Kenstowicz, Michael. 1996. Base Identity and Uniform Exponence: Alternatives to Cyclicity. In *Current Trends in Phonology: Models and Methods*, ed. Jacques Durand and Bernard Laks, 363–394. Manchester: European Studies Research Institute, University of Salford.
- Kenstowicz, Michael, and Hyang-Sook Sohn. 2001. Accentual Adaptation in North Kyungsang Korean. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 239–270. Cambridge, Mass.: MIT Press.
- Kim, Heung-gyu, and Beom-mo Kang. 2000. *Frequency Analysis of Korean Morpheme and Word Usage*. Seoul: Institute of Korean Culture, Korea University.
- Kiparsky, Paul. 1978. Analogical Change as a Problem for Linguistic Theory. Forum lecture, LSA Linguistic Institute, University of Illinois. Published in Kiparsky 1982, 217–236.
- Kiparsky, Paul. 1982. *Explanation in Phonology*. Dordrecht: Foris.
- Lee, Iksop, and Robert Ramsey. 2000. *The Korean Language*. Albany: State University of New York Press.
- Martin, Samuel E. 1954. *Korean Morphophonemics*. Baltimore: Linguistic Society of America.
- Martin, Samuel E. 1992. *A Reference Grammar of Korean*. Tokyo: Tuttle Publishing.

McCarthy, John J. 1991. Synchronic Rule Inversion. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 17: 192-207.

Sohn, Ho-Min. 1999. *The Korean Language*. Cambridge: Cambridge University Press.