Speakers have knowledge of many statistical generalizations about the distribution of linguistic forms, among them the allomorphs of non-automatic or morphophonological alternations.

There is more than one possibility, however, for how knowledge concerning allomorph distribution—patterns of alternation and non-alternation—is reflected in the phonological grammar:

- "Proportional Representation" (PR): patterns of alternation are phonologized as stochastic rules whose strength is proportional to the lexical frequency of the corresponding patterns.
- "Winner Take All" (WTA): only the most frequent pattern of alternation is phonologized; others ("minority alternations") are treated as irregular and subject to gradual elimination.

PR and WTA make distinct predictions about variation and its absence in innovative forms and the existing vocabulary:

In PR theories (e.g. Zuraw 2000), stochastic rules/constraints that reflect lexical statistics dictate the treatment of innovative forms. Established forms are held to have fixed, lexicalized pronunciations, so that the stochastic grammar must be prevented from applying to them. In sum, PR predicts:

- Innovative forms: variation according to lexical statistics
- Established forms: no variation

Under WTA, patterns of alternation are either regular or irregular. Regular patterns other than the null alternation are represented as rules, irregular patterns as excess information\(^1\) (e.g. listed allomorphs) in lexical entries. Assuming that such excess information tends to be lost over time, WTA predicts that items displaying irregular alternations will tend to vary with regularly derived substitutes. Innovative forms will normally lack excess information in their lexical entries and will therefore fail to show variation. WTA thus predicts:

- Innovative forms: no variation
- Established forms: variation if and only if irregular

**A test case**

Historically, Korean noun stems end in a variety of consonants and clusters; these appear before V-initial suffixes, but are subject in codas to cluster reduction and neutralization of laryngeal and manner contrasts, so that the only permissible coda obstruents are \(/p\ t\ k/\). With the reanalysis of isolation forms as basic, \(/p\ t\ k/\) have come to alternate unpredictably, when prevocalic, with the original stem-final consonants and clusters as follows (stem-type counts from Kim and Kang 2000\(^2\)).

---

\(^1\) I.e. in excess of a simple pairing of sound and meaning.

\(^2\) These counts were obtained using the concordance program Geuljabi II as applied to the .txt files of the CD-ROM accompanying that work. They correspond closely to the counts reported by Albright (2008:171) except with regard to stems showing the 't ~ s' alternation, for which Albright’s count is higher; similar counts are reported by Jun (2010:149) (neither Albright nor Jun include alternations involving clusters).
A PR analysis of these alternations will postulate stochastic rules/constraints with reliability/confidence levels that to a first approximation reflect the lexical statistics above (for an analysis of this sort based on the minimal generalization model of Albright and Hayes 2003 and relying for the most part on speaker well-formedness judgements, see Jun 2010). On the WTA analysis, the regular alternations will be \( t \sim s \) for coronals and the null alternation for labials and velars. The only rule postulated will thus be \( t \rightarrow s / . \) (Ko 1989). Minority alternations will be represented in lexical entries, e.g. by means of proper inclusion precedence formalism, with an irregular alternant constituting a special case and the basic form a default.\(^4\)

**Data**

While internet search results cannot be taken as reliable indications of frequency of occurrence, search engines do function as sensitive detectors of variation. In the Korean case, they reveal not only variation associated with ongoing regularization (e.g. prevocalic \( mul^h-/mulip- \) "knee" \( puok^h-/puok- \) "kitchen" \( juc^h-/jus- \) "stick game") that is widely reported but makes no appearance in standard corpora, but also variation of types that to my knowledge have not been generally noted (e.g. prevocalic \( sepjoknjak^h- \) for expected \( sepjoknjak^{k^h} \) "dawn").

This sensitivity means that when search results show zero variation for the inflection of a given stem, we can conclude with some confidence that that stem is indeed of invariant inflection. In the absence of an appropriate corpus of spoken or unedited written Korean, I use internet search results as a test for the presence versus absence of inflectional variation.

Search results for obstruent-final loanwords (ten stems for each point of articulation; six clitics and three stem allomorph candidates for each stem) show that, as predicted by the WTA, there is virtually no variation in their inflection. The exception is that some stem-final obstruents, most frequently labials, show a weak tendency to display aspiration, with \( /p^h t^h k^h/ \) appearing in place of \( /p t k/\). I suggest that this phenomenon represents reinterpretation of the aspiration that accompanies the stops in question when they occur before epenthetic \( /i/\). In illustration, I present search results for \( mak'et(i) \) "market" (www.google.com, 15 May 2013).

---

\(^3\) Jun (2010:156), following Albright (2005:41), is well aware that the Korean situation appears to call for a WTA condition, in that only the strongest of a set of competing alternations tends to be extended. He is nevertheless at pains to argue for the productivity of minority alternations and in particular (Jun and Lee 2007) for their activity in the phonology of loanwords.

\(^4\) Multi-layered proper inclusion precedence formalism is capable of mirroring relatively complex patterns of variation, e.g. the distinction between \( ^c^d^/-\)stems that tend to assimilate to \( ^c^h^/-\)stem conjugation and \( ^c^d^/-\)stems that assimilate directly to the default \( t \sim s \) alternation.
If this interpretation is correct, there would seem to be no evidence from innovative stems that minority alternations are phonologically active.

For established stems, we observe the dichotomy predicted by WTA:

- Historical *p-/*s-/*k-stems show essentially invariant inflection.
- Stems ending historically in other consonants and in clusters tend to show variation between the historically expected alternation and the regular one (cf. "knee", "kitchen", "stick game" above); there is also a well-known relation of mutual influence between (some) *th-stems and *ch-stems.

The data thus seem to support the WTA analysis. And, in view of the recorded history of the language, any other result would be surprising: in noun stems, extension of the \( t \sim s \) alternation for coronals and leveling of all non-coronal alternations has been in progress for centuries (Martin 1992, Ito 2010), suggesting that the WTA analysis has been in force over that period.

A methodological lesson: reliance on speaker well-formedness judgements (Jun 2010, Jun and Lee 2007) may create the appearance of more variation than is actually attested and obscure categorical generalizations that are in force.

But what about cases like Tagalog nasal substitution (Zuraw 2000), where loanwords do seem to track lexical statistics?

- At a descriptive level, the most salient fact is that speakers appear to have made no categorical generalization.
- At an explanatory level, the most interesting question is why they should have failed to do so.
- Research is needed on the relative frequency of the two types of case and what distinguishes them.

### In closing

I have claimed that innovative Korean noun stems are all regular. But we might expect an innovative stem to be irregular if it were a compound with an irregular last member— if, that is, it inherited irregularity from a constituent morpheme.

Diphthongization in innovative Spanish verbs

- shows that inheritance of irregularity from a constituent morpheme is possible.
- suggests that this is in fact the only way for an innovative stem to be irregular.
Of 20 verbs with stem vowels /e o/ that occur only in the 20th-century portion of the Corpus del Español, all but *interferir* "interfere" and *sobrevolar* "overfly" are non-diphthongizing. The former inherits diphthongization from the root *-fer-* (all eight other verbs in *-ferir* are diphthongizing), the latter from the root/stem *vol*-. Non-diphthongizing *proliferar*; with stem vowel /e/ in the same environment as *interferir* underlines the fact that it is morphemic identity rather than phonological context that determines whether diphthongization will be extended.

If the only way for an innovative stem to display diphthongization is for it to inherit the alternation from a constituent morpheme, it would seem that the alternation must be a categorical property of particular lexical items.

**Where from here?**

The WTA theory, in assigning minority alternations to the lexicon rather than the grammar, brings into focus the distinction between cases in which speakers do generalize a non-automatic alternation and those in which they gradually eliminate it. Spanish Diphthongization (*e/o → ie/ue*), which has been leveled far more frequently than extended, and Portuguese Lowering (*e/o → ɛ/ɔ*), which has been fully generalized, constitute a suggestive minimal pair in this regard: given that the Portuguese alternants differ by only a single degree of vowel height, the difference between the two cases could be explained if there is an input-output phonological distance condition on the induction of morphophonological rules.

**References**


**Acknowledgements**

Many thanks to Sang-Cheol Ahn, Jae-Woong Choe, Sukhoon Choo, Mark Davies, Yoonjung Kang, and Yunjin Nam for invaluable assistance of various kinds. None of them are responsible for my mistakes.

contact: dechene@waseda.jp

---

5 www.corpusdelespanol.org. Three of the verbs occur marginally in the 19th-century portion of the corpus as well; verbs in *-ear* and *-ionar*, which never show diphthongization, were eliminated from the candidate pool.