

# Lexical exceptionality in paradigm-specific learning: modeling stem-final obstruent alternations in Korean verbs and adjectives

Stella Eunsoo Hong  
University of Utah  
u1597092@utah.edu

## Abstract

Korean stem-final conjugations illustrate the interaction between lexical exceptionality and heterogeneous phonological processes. When /p/-, /t/-, and /s/-final stems occur before vowel-initial suffixes, the irregular classes in these paradigms undergo intervocalic lenition, each exhibiting a distinct alternation pattern. Learners must therefore not only identify which roots trigger lenition, but also determine the corresponding repair strategy. This study investigates how lexically-specific phonological patterns are acquired when multiple repair strategies are available. We employ a lexically scaled MaxEnt model (Linzen et al., 2013; Hughto et al., 2019) to learn these paradigm-specific alternations and run simulations under two learning scenarios: (1) when repair strategies occur at equal frequencies and (2) when one strategy significantly outnumbers the others. Results show that the model favors a *least-cost* solution by treating statistically dominant morpheme classes as the general pattern. We conclude by discussing the model's sensitivity to lexical statistics, predictions for empirical testing, and implications for language acquisition.

## 1 Introduction

Lexical exceptionality requires learners to reconcile competing generalizations within a single grammatical system. When exceptional morphemes resist the general pattern, they challenge the constraint rankings that otherwise account for the majority of forms. Within Optimality Theory (OT), two core approaches to modeling such conflicts are cophonology (Inkelas et al., 1997; Anttila, 2002; Inkelas and Zoll, 2003) and lexical indexation (Pater, 2000; Coetzee, 2009). Cophonology assumes separate rankings for different morphological classes, thereby permitting multiple subgrammars to resolve the conflict. Lexical indexation, by contrast, indexes constraints to specific lexical classes, allowing a single ranking to accommodate

the conflict through lexically restricted application of the constraint.

Previous literature has largely focused on exceptionality involving a single phonological process. Less attention has been paid to cases in which multiple repair processes coexist, which can potentially introduce substantial ranking complexity under both approaches introduced above. Nevertheless, many languages employ more than one strategy to repair ill-formed phonological sequences (Pater, 1999). When such competing repair strategies intersect with lexical exceptionality, learners must determine not only how exceptional and non-exceptional items are represented, but also how competing repairs are ranked against one another. This raises the question of how learners reconcile inconsistency across two dimensions: lexical specificity and competing repair strategies.

Korean stem-final conjugations exemplify such empirical patterns and serve as a useful test case for this question. When irregular classes of verbs and adjectives ending in /p/, /t/, and /s/ combine with vowel-initial suffixes, the stem-final obstruents surface as [w], [r], and  $\emptyset$ , respectively, through intervocalic lenition. All three lenition strategies are motivated by avoidance of the dispreferred \*V-OBS-V sequence, but each applies only to a specific stem-final paradigm.

To address the learning of *paradigm-specific lexical exceptionality*, we adopt a lexically scaled MaxEnt framework (Linzen et al., 2013; Hughto et al., 2019). As a probabilistic implementation of the indexation approach, the model encodes both lexical exceptionality and competing repair strategies within a single set of weights. Lexical specificity is implemented through additive morpheme-specific weights associated with particular paradigm classes, thereby capturing the role of indexed constraints while avoiding the need to posit separate subgrammars.

Moreover, the distribution of irregular forms in

Korean varies across stem classes, allowing us to examine how statistical properties influence the choice of repair strategy. The majority of /p/-final stems exhibit irregular alternations, whereas /t/- and /s/-final stems show roughly equal proportions of regular and irregular forms. To investigate the effect of this distributional asymmetry, we ran a toy simulation under two learning scenarios: (1) when only /t/- and /s/-final paradigms with roughly equal proportions of regular and irregular forms are considered, and (2) when the /p/-final paradigm, which exhibits a strong majority of irregular forms, is added. For both experimental conditions, the number of suffixes was incrementally increased to explore how the degree of affix exposure influences success in learning stem-specific exceptionality.

Results show improved generalization as the number of suffix types increases.<sup>1</sup> When given only a few suffix types, additive morpheme-specific weights were incorrectly assigned to suffixes. As the suffix inventory grew, these weights gradually shifted toward stems, correctly identifying them as the locus of lexical exceptionality. This process first involved assigning scaling weights selectively to irregular classes, thereby distinguishing them from regular classes, and was later refined as the model began to capture paradigm-specific distinctions among irregular forms. In the experiment including the /p/-final paradigm, the learner treated the /p/-irregular alternation as the default repair when lenition was triggered, as the constraint violated by this class showed the greatest resistance to receiving general weight. This configuration came at the cost of reduced prediction accuracy for the other two alternating classes.

The rest of the paper is organized as follows. Section 2 introduces the empirical patterns. Section 3 reviews the relevant literature and motivates the use of lexically scaled MaxEnt to model Korean stem-final conjugation. Section 4 presents the implementation details of the simulations, and Section 5 summarizes the results. Finally, Section 6 discusses the implications of these findings and outlines how learning may proceed when real-world data are fully incorporated.

<sup>1</sup>Our experimental code and data are available at [https://github.com/stxllastar/scaled\\_conjugation](https://github.com/stxllastar/scaled_conjugation).

## 2 Korean stem-final verb / adjective conjugations

Korean verbs and adjectives are traditionally divided into regular and irregular classes. Regularity is defined in terms of form invariance: regular forms remain unchanged within a paradigm, whereas irregular forms exhibit alternation before certain suffixes (Lee, 1996; Rhee, 1996). For clarity, we refer to irregular and regular classes as alternating and non-alternating classes, respectively<sup>2</sup>.

Alternating conjugation patterns can be further classified into three types depending on the locus of alternation: (1) the stem, (2) the suffix (i.e., the inflectional ending), or (3) both. Within the first type, /p/-, /t/-, and /s/-final stems constitute a particularly interesting case, as their alternations involve intervocalic lenition conditioned by the same phonological environment.

(1) Intervocalic lenition in alternating classes

- a. p-stems: /p/ → [w] / V\_V
- b. t-stems: /t/ → [r] / V\_V
- c. s-stems: /s/ → ∅ / V\_V

As (1) illustrates, when vowel-initial suffixes follow alternating-class stems, the stem-final obstruents /p/, /t/, and /s/ undergo lenition, surfacing as the sonorants [w] and [r] or deleting (Lee, 2013; Albright and Kang, 2009). Because Korean syllable structure disallows complex onsets and codas, vowel-initial suffixation places stem-final obstruents in an intervocalic environment. This lenition process is lexically restricted, as stems belonging to the non-alternating class remain invariant in the same environment—a contrast illustrated in Table 1.

Class	Stem	Conjugation
p-alternating	/ka.kap/ 'close'	[ka.ka.w-Δ]
p-non-alternating	/ip/ 'wear'	[i.p-Δ]
t-alternating	/kje.dat/ 'realize'	[kje.da.r-a.jo]
t-non-alternating	/mut/ 'bury'	[mu.t-Δ]
s-alternating	/nas/ 'better'	[na.-a.do]
s-non-alternating	/us/ 'laugh'	[u.s-Δ]

Table 1: Alternating vs Non-alternating classes in V\_V

Meanwhile, when stems are followed by consonant-initial suffixes, the environment triggering alternation is absent. Consequently, the contrast between classes observed in Table 1 disappears: the

<sup>2</sup>The traditional distinction between regular and irregular classes does not directly correspond to phonological regularity, as the forms labeled *irregular* undergo systematic alternations and in fact outnumber the regular forms.

underlying /p/, /t/, and /s/<sup>3</sup> surface faithfully in both classes. As shown in Table 2, both alternating and non-alternating stems preserve their underlying codas before the declarative suffix *-ta*.

Class	Stem	Conjugation
p-alternating	/ka.kap/ 'close'	[ka.kap.-ta]
p-non-alternating	/ip/ 'wear'	[ip.-ta]
t-alternating	/kje.dat/ 'realize'	[kje.dat.-ta]
t-non-alternating	/mut/ 'bury'	[mut.-ta]
s-alternating	/nas/ 'better'	[nas.-ta]
s-non-alternating	/us/ 'laugh'	[us.-ta]

Table 2: Alternating vs Non-alternating classes in V\_C

To summarize, the phenomenon exhibits three key properties. First, it shows lexical idiosyncrasy: only alternating (irregular) stems undergo lenition. Second, it involves paradigm-specific repair, whereby each obstruent-final alternating class adopts a distinct lenition strategy to avoid the \*V-OBS-V configuration. Third, the process reflects stem-based exceptionality, with suffixes merely providing the phonological environment that triggers the alternation (intervocalic with -V suffixes and non-intervocalic with -CV suffixes). Hereafter, we refer to suffixes as *environmental triggers*.

### 3 Previous approaches towards learning lexical specificity

#### 3.1 The lexical indexation approach

Among the two widely discussed approaches to lexical exceptionality—cophonology and lexical indexation—Korean conjugation has been analyzed within the framework of lexical indexation. Lee (2013) adopts the Recursive Constraint Demotion (RCD) algorithm (Tesar and Smolensky, 1998; Tesar and Prince, 2003) in conjunction with constraint cloning (Pater, 2009) to address the ranking conflict between alternating and non-alternating classes within each paradigm.

RCD incrementally constructs a constraint hierarchy by installing the best-performing constraint—one that prefers only winning candidates—immediately below the established ranking. However, in cases of lexical exceptionality, no single constraint ordering satisfies all winner–loser comparisons: ranking \*V-OBS-V below

<sup>3</sup>When /s/-final stems are followed by consonants, the coda surfaces as [t] due to an independent coda neutralization rule, mapping all obstruents to their homorganic lenis stop in syllable-final position (Jun, 2010). This rule is not relevant to the alternation under investigation and is therefore set aside in the present analysis.

the faithfulness constraint correctly favors the non-alternating-class winner but incorrectly prefers the losing candidate for the alternating class, whereas ranking \*V-OBS-V above the faithfulness constraint produces the opposite result.

The conflicting requirements for alternating stems (\*V-OBS-V  $\gg$  FAITH) and non-alternating stems (FAITH  $\gg$  \*V-OBS-V) are resolved by cloning \*V-OBS-V and indexing its applicability separately to the two classes. This yields the ranking scheme in (1), where both orderings are satisfied through a partial ordering of class-specific constraints.

$$*V-OBS-V_{Alter} \gg FAITH \gg *V-OBS-V_{Non-Alter} \quad (1)$$

What remains unclear is how separate instantiations of (1) for individual paradigms collectively account for the grammar as a whole. Because candidates in the /p/-, /t/-, and /s/-paradigms adopt different repair strategies and violate distinct faithfulness constraints, competition arises among the faithfulness constraints occupying the ranking space between \*V-OBS-V<sub>Alter</sub> and \*V-OBS-V<sub>Non-Alter</sub>.

In the /p/-paradigm, the relevant faithfulness constraint violated by repair to [w] is IDENT(CONSONANTAL); in the /t/-paradigm, repair to [r] violates IDENT(LATERAL); and in the /s/-paradigm, deletion of the obstruent segment violates MAX-C. Consequently, in each paradigm, the faithfulness constraint implicated in the repair must be outranked by the competing faithfulness constraints associated with the other two paradigms. To capture these cross-repair ranking relations, the ranking scheme in (1) must be expanded as in (2).

$$\begin{aligned} &*V-OBS-V_{P-Alter} \gg IDENT(LATERAL), MAX-C \gg \\ &IDENT(CONSONANTAL) \gg *V-OBS-V_{P-Non-Alter} \\ &*V-OBS-V_{T-Alter} \gg IDENT(CONSONANTAL), MAX-C \gg \\ &IDENT(LATERAL) \gg *V-OBS-V_{T-Non-Alter} \\ &*V-OBS-V_{S-Alter} \gg IDENT(CONSONANTAL), \\ &IDENT(LATERAL) \gg MAX-C \gg *V-OBS-V_{S-Non-Alter} \end{aligned} \quad (2)$$

Undesirably, the expanded ranking requires the learner to maintain paradigm-specific instantiations of the grammar in addition to indexed constraints. Consequently, the learner must track both paradigm-specific rankings and the application domains of cloned constraints, thereby inheriting the learning burdens of both lexical indexation and the cophonology approach.

### 3.2 Lexically Scaled MaxEnt

To address this challenge in a more tractable setting, the present study adopts a probabilistic implementation of indexation—the lexically scaled MaxEnt model of Linzen et al. (2013); Hughto et al. (2019)—which captures paradigm-specific patterns without positing separate ranking systems. The model is a lexicalized variant of Maximum Entropy (MaxEnt) grammar (Goldwater and Johnson, 2003), designed to account for languages exhibiting both variability and lexical exceptionality. Non-lexically-specific variability is captured through general constraint weights, whereas lexically idiosyncratic patterns are modeled through morpheme-specific scaling factors incorporated into the harmony score used to compute MaxEnt probabilities.

$$\mathcal{H}_i = \sum_{\gamma \in \Gamma} \left( w_\gamma + \sum_{m \in \mu_i} s_{\gamma m} \right) v_{\gamma i} \quad (3)$$

In (3), morpheme-specific scaling factors  $s_{\gamma m}$  are added to the general constraint weight  $w_\gamma$  before being multiplied by candidate  $i$ 's violation profile,  $v_{\gamma i}$ , over the set of constraints  $\gamma \in \Gamma$ . As indicated by the subscripts, the general weight  $w_\gamma$  applies grammar-wide, whereas the scaling factors  $s_{\gamma m}$  are specific to the individual morphemes  $m$  that constitute a given candidate. This additive penalty is conceptually analogous to lexical indexation: a higher-ranked class-specific markedness constraint (e.g., \*V-OBS-V<sub>Alter</sub> in (1)) can be recast as an increased scaling factor applying to that lexical class. To illustrate this correspondence, consider the idealized learning outcome for the present phenomena shown in Table 3.

	*V-OBS-V	MAX	ID-LAT	ID-C
General weights	4	6	6	6
p-alternating	6	3	2	0
p-non-alternating	0	1	1	1
t-alternating	6	3	0	4
t-non-alternating	0	1	1	1
s-alternating	6	0	3	3
s-non-alternating	0	1	1	1
suffixes	0	0	0	0

Table 3: Hypothetical Learning Result

The rows beneath the general weights indicate scaling factors associated with each morpheme class, from which three learning predictions follow. First, \*V-OBS-V must outrank the faithfulness constraints in alternating classes but not

in non-alternating classes. Accordingly, only alternating classes bear a positive scaling weight of 6 on \*V-OBS-V. As larger weights translate into higher rankings, alternating classes achieve the ordering \*V-OBS-V<sub>Alter</sub>  $\gg$  FAITH, whereas non-alternating classes, receiving no scaling, obey the reverse ordering FAITH  $\gg$  \*V-OBS-V<sub>Non-Alter</sub>. These two partial orders constitute (1).

Second, for the correct repair strategy to emerge within each paradigm, the faithfulness constraint violated by the paradigm-specific alternation must receive the lowest scaling weight: ID-C for the /p/-alternation, ID-LAT for the /t/-alternation, and MAX for the /s/-alternation<sup>4</sup>. In Table 3, this is implemented by assigning a scaling weight of 0 to the violated faithfulness constraint in each alternating class, producing the effect of cross-repair competition within the paradigm.

Third, suffixes that function as *environmental triggers* ideally bear no scaling weights, since they are not themselves the locus of exceptionality.

When the learned grammar evaluates each conjugated item, the scaling factors associated with its constituent morphemes combine with the general weights to determine the effective constraint ranking. A non-lenited realization /ka.kap- $\Lambda$ / in a p-alternating class (Table 1), for example, would be calculated as 4 + 6 + 0, consisting of the general weight, stem-scale, and suffix-scale. In this way, distinctions along two dimensions—paradigm and exceptionality—are captured while avoiding the ranking complexity posited by the expanded indexation approach in (2).

Among the four hypothetical languages explored in Hughto et al. (2019), the *lexical language* most closely parallels the present phenomenon. In that language, lexically indexed prefixes categorically trigger vowel deletion, analogous to how different stem classes categorically opt for different strategies. In their learned grammar, deletion-triggering and deletion-blocking prefix classes receive complementary scaling on ALIGN and \*CCC, such that the violated constraint in each class receives the smaller scaling weight (Table 4). This mirrors the first and second predictions proposed here, where non-alternating classes receive no scaling on \*V-OBS-V, while the constraint violated by a paradigm-specific repair receives the lowest scaling within each alternating class. Moreover, because

<sup>4</sup>For notational brevity, this modeling work abbreviates IDENT(CONSONANTAL), IDENT(LATERAL), and MAX-C as ID-C, ID-LAT, and MAX.

stems are not the locus of lexical specificity, they receive no scaling, paralleling the present analysis in which suffixes ideally bear no scaling weights.

	*CCC	MAX	ALIGN
General Weights	4.6	0.0	4.1
Deleting Prefix	0.0	0.0	6.4
Non-Deleting Prefixes	5.3	0.0	0.0
Stems	0.0	0.0	0.0

Table 4: Learned weights of the *lexical language* (Hughto et al., 2019)

## 4 Experiments

### 4.1 Toy data Generation

Table 5 reproduces the statistics on Korean obstruent-final stems from the Sejong Corpus reported in Lee (2013). The Sejong Corpus, compiled by the National Institute of Korean Language (NIKL) as part of the 21st Century Sejong Project, has been widely used in previous studies (Albright and Kang, 2009; Jun, 2010), making it a well-established source for such statistics.

p-alternating	68 (72.34%)
p-non-alternating	5 (5.32%)
t-alternating	5 (5.32%)
t-non-alternating	7 (7.45%)
s-alternating	5 (5.32%)
s-non-alternating	4 (4.26%)

Table 5: Distributions of /p/-, /t/-, and /s/-final stems in the Sejong Corpus

Following the type counts in Table 5, we generated pseudo-stems with a CV-/p,s,t/ structure, reflecting Korean syllable phonotactics that restrict consonant clusters. Using the Romanization system provided by NIKL<sup>5</sup>, onset consonants and vowels were drawn from the Romanized inventory of Korean segments. Onsets were selected from {g, kk, k, d, tt, t, b, pp, p, j, jj, ch, s, ss, h, n, m, r}, while vowels were restricted to the set of monophthongs {a, eo, o, u, eu, i, ae, e, oe, wi}. Random sampling from these sets yielded 180 unique CV sequences. These sequences were assigned without overlap to the six conjugation classes listed above and instantiated with the class-appropriate stem-final consonant (/p/, /t/, or /s/). The full list of toy stems is provided in Table 12 in Appendix A.

Because alternation is triggered in an intervocalic context, the inventory of novel suffixes was

<sup>5</sup>[https://www.korean.go.kr/front\\_eng/roman/roman\\_01.do](https://www.korean.go.kr/front_eng/roman/roman_01.do)

balanced with respect to onset presence: suffixes beginning with a consonant (–CV endings) and those beginning with a vowel (–V endings) were equally represented. The initial inventory consisted of two minimal endings, –ta and –a. It was then incrementally expanded by introducing additional vowel-initial and /t/-initial pairs, with vowels drawn from the same monophthong set used in stem generation. Table 13 in Appendix A summarizes the suffix sets across experimental conditions.

### 4.2 Distributional Conditions

Table 5 shows that /p/-alternating stems constitute the overwhelming majority of the data (72.34%). To isolate the effect of this distributional skew on the learner, we first conducted experiments under a simplified setting that excluded the /p/-paradigm.

The modeling data were initially restricted to /t/- and /s/-final stem conjugations, which we refer to as the *balanced condition*. In this setting, we examine two repair strategies for resolving \*V-OBS-V: deletion and acquisition of the [+lateral] feature. Accordingly, triplets of UR–SR mappings are considered for each input form, corresponding to (a) the faithful candidate, (b) lenition via violation of IDENT(LATERAL), and (c) lenition via violation of MAX-C. For example, a t-alternating stem /kjɛ.dat/ in Table 1 would yield the candidates (a) [kjɛ.dat-a], (b) [kjɛ.dal-a], and (c) [kjɛ.da-a]. Table 14 in Appendix B presents the UR–SR mappings for the four stem classes under the two suffix conditions introduced above (–V and –CV endings).

Subsequently, we extended the model to the full paradigm by incorporating the /p/-ending data, creating the *imbalanced condition*. In this setting, loss of the [+consonantal] feature becomes an additional repair strategy alongside deletion and acquisition of [+lateral]. Accordingly, UR–SR mappings are represented as quadruplets for each input form, corresponding to (a) the faithful candidate, (b) lenition via violation of IDENT(CONSONANTAL)<sup>6</sup>, (c) lenition via violation of IDENT(LATERAL), and (d) lenition via violation of MAX-C. Extending the previous example, /kjɛ.dat/ would now additionally yield (b) [kjɛ.daj-a] in addition to the previous three candidates. Table 15 in Appendix B presents a subset of the input tableaux under the two suffix

<sup>6</sup>When constructing candidate (b) for the /t/- and /s/-paradigms, the segmental alternation was instantiated as [j], a glide that preserves the place of articulation, similar to how /p/ surfaces as [w] while preserving labiality.

conditions.

### 4.3 Model Parameters

Learner settings followed the default implementation in Hughto et al. (2019). We used an L1 prior ( $\lambda = 1.0$ ) and optimized the model using gradient descent with clipping (Tsuruoka et al., 2009). Weights and scales were initialized to 0.0, and training proceeded for 20,000 epochs with a learning rate of 0.001. All parameters were held constant across experimental conditions.

## 5 Results

### 5.1 Learning with the /t/- and /s/- paradigms

When only /t/- and /s/-final stems were included, type frequencies were roughly balanced both across paradigms and within each paradigm (Table 5). Nevertheless, the attested forms exhibit a 1:3 alternation ratio, because intervocalic lenition occurs only when alternating stems are followed by vowel-initial suffixes; in the other three configurations—alternating stem + CV suffix, non-alternating stem + V suffix, and non-alternating stem + CV suffix—faithful realization prevails. Given this overall distribution, the learner assigned less general weight to \*V-OBS-V than to IDENT(LATERAL) and MAX-C.

	*V-OBS-V	MAX	ID-LAT
General weights	0	0.79	0.79
t-alternating stems	0	0	0
t-non-alternating stems	0	0	0
s-alternating stems	0	0	0
s-non-alternating stems	0	0	0
-V suffix	0	0	0
-CV suffix	0	2.16	2.16

Table 6: Learned weights (/t, s/ stems; two suffixes)

As such, while the majority pattern—faithful surfacing—is captured through general weights, the alternations are handled through lexical scaling. This division reflects the model’s *least-cost* principle, which prioritizes general mechanisms for capturing dominant patterns while reserving lexical mechanisms for more restricted ones. The same pressure toward economy also shapes the learning of scaling factors. Jarosz (2025) shows that lexicalization tends to be stronger for higher-frequency items than for lower-frequency ones, and this tendency manifests in scaling weights initially being misattributed to the non-idiosyncratic suffixes.

When only two suffix types are available (Table 6), each stem occurs only twice in the attested forms—once with *-ta* and once with *-a*—while these two suffixes combine with all stem types. Consequently, when frequencies are evaluated per morpheme, suffixes occur more often than stems and thus emerge as the more general morpheme selected to account for exceptionality.

The same frequency bias persists within the suffix inventory. Because intervocalic lenition occurs in only one-fourth of the training data, the learner favors the -CV suffix, which creates a non-intervocalic environment, over the -V suffix. As a result, the highly frequent faithful realizations associated with the -CV suffix are preferentially captured through lexical scaling on MAX and ID-LAT rather than through scaling on \*V-OBS-V in the intervocalic (-V suffix) environment.

	*V-OBS-V	MAX	ID-LAT
General weights	2.2	3.26	3.26
t-alternating stems	0.37	0	0
t-non-alternating stems	0	0	0
s-alternating stems	0.37	0	0
s-non-alternating stems	0	0	0
-V suffix	0	0	0
-CV suffix	0	0	0

Table 7: Learned weights (/t, s/ stems; four suffixes)

As the number of suffixes increases, stems combine with a larger set of suffixes and occur more frequently in the attested forms. This increases the likelihood that stem morphemes are lexicalized, reflected in the progressive redistribution of scale weights from suffixes to stems in Tables 7 and 8, and in Tables 16 and 17 in Appendix C.

The redistribution proceeds in stages. First, the learner captures the triggering of alternation by distinguishing alternating and non-alternating classes, assigning scale weight to \*V-OBS-V in the former but not in the latter (Table 7). It then learns the paradigm-specific repair, establishing the ranking among faithfulness constraints that compete to derive class-specific forms. From Table 8 onward, the /t/- and /s/-alternating classes assign less scale weight to the violated constraint—ID-LAT and MAX, respectively—than to the competing faithfulness constraint, and this gap widens as additional suffix types are introduced.

This progression is expected under a frequency-sensitive learning mechanism. The /t/- and /s/-alternating classes have identical type frequencies (five types each), so the statistical signal for learn-

ing paradigm-specific behavior comes from the corresponding non-alternating classes. These differ only slightly in type frequency (seven for /t/-non-alternating and four for /s/-non-alternating), requiring more extensive exposure to attested forms before learners can acquire paradigm-specific distinctions beyond simply learning when alternation is triggered. In the present case, four suffix types are minimally required to distinguish alternating from non-alternating classes, and six suffix types to differentiate class-specific alternations.

	*V-OBS-V	MAX	ID-LAT
General weights	1.99	3.66	3.66
t-alternating stems	1.73	0.13	0
t-non-alternating stems	0	0	0
s-alternating stems	1.73	0	0.13
s-non-alternating stems	0	0	0
-V suffix	0	0	0
-CV suffix	0	0	0

Table 8: Learned weights (/t, s/ stems; six suffixes)

Improvements in the distribution of scale weights are reflected in the model’s predictive performance. To evaluate this effect, we focus on alternating stems in intervocalic contexts, which represent the minority lenition-triggering portion of the training data and are thus harder for the learner to acquire. As shown in Fig. 1, accuracy in predicting the correct repair increases with the number of suffix types<sup>7</sup>. Notably, accuracy exceeds the 33% threshold only in the six-suffix condition<sup>8</sup>, where the number of suffix types matches or exceeds the number of non-alternating stem types per paradigm (7 for the /t/-paradigm and 4 for the /s/-paradigm).

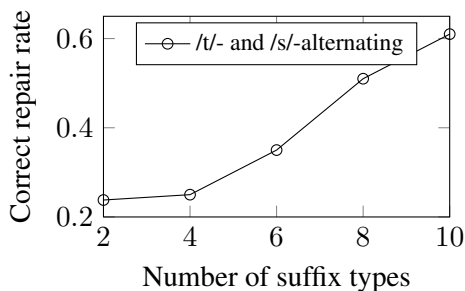


Figure 1: Repair accuracy of /t/- and /s/-alternating stems in intervocalic contexts

<sup>7</sup>The full numerical results for training accuracy are provided in Table 20 in Appendix D.

<sup>8</sup>Chance performance with three possible SR forms is approximately 33%.

## 5.2 Learning with the /p/, -/t/, and -/s/ paradigms

Accommodating the /p/-final data creates an imbalance among the three alternating patterns. As shown in Table 5, the /p/-alternating class (72.34%) overwhelmingly outnumbers the /t/-alternating (5.32%) and /s/-alternating classes (5.32%). Although environments triggering intervocalic lenition remain outnumbered by those blocking it (approximately in a 1:3 ratio), most lenition tokens come from the /p/-alternating class, such that the number of glide-forming lenition tokens (68) approaches the frequency of faithful realizations (130).<sup>9</sup> Under these conditions, violation of ID-C emerges as the second default pattern after faithful realization. Consequently, when intervocalic lenition is triggered in the -V context, alternating stems favor repair via violation of ID-C.

This skewed distribution directly shapes the model’s learning behavior. First, across all suffix conditions (Tables 9–11 and Tables 18 and 19 in Appendix C), the /p/-alternating class never acquires stem-scale weights. Second, ID-C, the constraint violated by this class, shows the greatest resistance to receiving general weight. In the initial two-suffix condition, all of its weight is borne by the *environmental trigger* -CV suffix. This bias persists as the suffix inventory increases and continues to shape the model’s learning trajectory.

	*V-OBS-V	MAX	ID-LAT	ID-C
General Weights	0.69	2.6	2.6	0
p-alternating stems	0	0	0	0
p-non-alternating stems	0	0	0	0
t-alternating stems	0	0	0	0
t-non-alternating stems	0	0	0	0
s-alternating stems	0	0	0	0
s-non-alternating stems	0	0	0	0
-V suffix	0.69	0	0	0
-CV suffix	0	1.92	1.92	4.51

Table 9: Learned Weights (/p, t, s/ stems; two suffixes)

Excluding the effect of ID-C’s resistance to receiving general weight, the progression from Tables 9–19 largely parallels the learning results obtained when the /p/ paradigm is excluded. With only two suffixes available, stems initially lack sufficient morphological evidence to learn their indexed status. The second stage—when the model begins to differentiate alternating from non-alternating classes (here limited to /t/ and /s/, since

<sup>9</sup>Comprised of all non-alternating stems in both suffix conditions ((5 + 7 + 4) × 2) together with alternating stems in the non-triggering (-CV) environment (68 + 5 + 5).

/p/ no longer constitutes the minority pattern captured by lexical scale)—is delayed by the dominance of the /p/-alternations.

	*V-OBS-V	MAX	ID-LAT	ID-C
General Weights	4.02	5.21	5.21	2.14
p-alternating stems	0	0	0	0
p-non-alternating stems	0	0	0	1.41
t-alternating stems	0	0	0	1.41
t-non-alternating stems	0	0	0	1.41
s-alternating stems	0	0	0	1.41
s-non-alternating stems	0	0	0	1.41
-V suffix	0	0	0	0
-CV suffix	0	0	0	2.15

Table 10: Learned Weights (/p, t, s/ stems; four suffixes)

This is because alternation into a glide functions as a second default pattern alongside faithful surfacing, leaving the /t/- and /s/-alternating classes to compete with two majority patterns. Consequently, even in the four-suffix condition (Table 10), the model fails to differentiate among the six stem classes, as their stored vectors remain identical. Differentiation emerges only in Table 11, where scale weights begin to diverge between the /t/- and /s/-alternating classes.

	*V-OBS-V	MAX	ID-LAT	ID-C
General Weights	4.98	6.29	6.29	2.71
p-alternating stems	0	0	0	0
p-non-alternating stems	0	0	0	2.54
t-alternating stems	0.61	0	0	2.88
t-non-alternating stems	0	0	0	2.54
s-alternating stems	0.61	0	0	2.88
s-non-alternating stems	0	0	0	2.54
-V suffix	0	0	0	0
-CV suffix	0	0	0	1.53

Table 11: Learned Weights (/p, t, s/ stems; six suffixes)

The effect of statistical imbalance is further reflected in prediction accuracy across the alternating classes. As shown in Fig. 2, the curves for the /t/- and /s/-alternations fail to exceed the 25% chance-level threshold even in the six-suffix condition, unlike in the condition without /p/-alternation<sup>10</sup>. In contrast, the /p/-alternation maintains substantially higher accuracy throughout training. Nevertheless, the gap between the /p/-paradigm and the /t/- and /s/-paradigms gradually narrows as the number of suffix types increases, again highlighting the importance of suffix-type diversity.

<sup>10</sup>Chance-level accuracy in the imbalanced condition is 25%, since each UR corresponds to four possible SR candidates. See Table 21 in Appendix D for the exact numerical values.

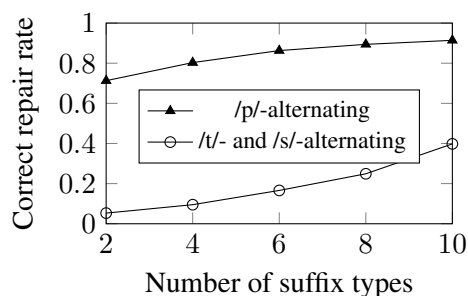


Figure 2: Repair accuracy of /p/-, /t/-, and /s/-alternating stems in intervocalic contexts

## 6 Discussion

### 6.1 Lexical statistics

One of the main findings is that, without a sufficiently large number of attested morphemic combinations, generalization cannot develop at the locus of exceptionality. This result highlights the statistical sensitivity of the model when learners are not exposed to enough morphological contexts in which exceptional items may occur.

The success of assigning scaling weights to the lexically idiosyncratic morphemes in the *lexical language* (see Table 4) can be interpreted in the same light. In Hughto et al. (2019), indexed items (prefixes) were outnumbered by non-indexed ones (stems), and the latter were replicated tenfold to achieve the reported performance. By contrast, our experiments involved the opposite asymmetry, where indexed items outnumbered non-indexed ones, and this unfavorable distribution was not compensated for through token-frequency amplification. Comparing the two results suggests that learning the locus of exceptionality critically depends on sufficient attestation in the input.

This observation raises the question of how suffixes are distributed in natural language. As non-indexed items, suffix statistics have received little attention and lack systematic documentation. We therefore conducted a corpus analysis of the Korean Google Treebank<sup>11</sup>. The corpus shows a distribution of stem types analogous to that in Table 5, while the suffix inventory far exceeds ten, comprising 29 -CV and 33 -V suffixes. Given such suffix type diversity in the empirical corpus data, absolute type frequency seems unlikely to pose a serious obstacle for real-world language acquisition.

<sup>11</sup><https://github.com/emorynlp/ud-korean/tree/master/google>. The stem and suffix inventories identified in the corpus are reported in Tables 23 and 24 in Appendix E

## 6.2 Empirical Simulation

Nevertheless, an additional finding from our corpus analysis was that, although the number of available suffix types is sufficiently large, stems are not attested with every suffix type. In our earlier idealized learning scenario, full paradigm saturation was assumed, whereby every stem combines with all possible suffixes (Chan, 2008; Kodner, 2022). This assumption poorly reflects the sparsity of naturalistic input data, which typically follows a heavily skewed Zipfian distribution: a small number of stems occur across many derivational forms, whereas most combine with only a few suffixes.

This mismatch raises the possibility that learning difficulty may arise not from absolute type frequency, but from limited paradigm saturation. To test this hypothesis, we conducted an additional simulation using conjugation patterns extracted from the corpus. The results revealed the effects of Zipfian skewness: despite the availability of sufficient suffix types, stems still struggled to acquire lexical scaling, with only a small subset receiving nonzero scaling weights and even fewer learning sizable values.

By contrast, sizable scaling factors were widely assigned to suffixes, likely reflecting their broader attestation across stems. While most stems are attested with only a few suffixes, suffixes themselves appear with many different stems. This situation parallels the earlier two-suffix condition, in which stems were attested only twice; here, however, the challenge arises not from a limited number of morphemic counterparts, but from limited paradigm saturation across stem–suffix combinations.

Taken together, these results indicate that, under naturalistic input conditions, learning difficulty lies not in the absolute number of available types, but in the asymmetric degree of paradigm saturation between stems and suffixes. Furthermore, the Google Treebank UD corpus used in the empirical simulation has been identified as a suitable non-CDS proxy for acquisition input in computational learning systems (Kodner, 2022), supporting the ecological validity of the present findings.

## 6.3 Language acquisition

While analytically unexpected, the initial two-suffix condition—where lexical idiosyncrasy is attributed to suffixes—may be interpreted as a stage of overregularization. Because suffixes provide contextual information and function as phonologi-

cally regular *environmental triggers*, learners may initially assign excessive weight to their behavior, associating alternation with a general phonological rule before learning the locus of exceptionality itself. In particular, lenition is initially conditioned by suffixal phonological context rather than by stem-specific lexical exceptionality: it occurs before vowel-initial suffixes but not consonant-initial suffixes (Tables 6 and 9).

Such an interpretation may further address the critical discrepancy between model behavior and human language acquisition identified in Jarosz (2025), wherein regularization effects are strongest at early stages of learning in humans but are predicted to consistently strengthen over the course of training in Scaled MaxEnt.<sup>12</sup> Increasing the suffix inventory approximates a learning scenario in which the learner is exposed to progressively richer morphological exponence. Under this interpretation, the two-suffix condition corresponds to an early stage in the model’s learning trajectory, where alternation is generalized as a phonologically conditioned pattern before being refined into stem-specific lexical exceptionality. As observed throughout Tables 7, 8, 16, and 17, and Tables 10, 11, 18, and 19, increasing the suffix inventory progressively reduces this tendency toward regularization, suggesting that the model’s behavior may not entirely diverge from developmental patterns observed in human acquisition.

Nevertheless, two gaps remain if the model is to fully account for human learning of paradigm-specific exceptionality. First, it must address how appropriate scaling factors can be learned under conditions of limited paradigm saturation across stem–suffix combinations. Second, it must account for how statistically non-dominant repair strategies—such as those in the /t/ and /s/ paradigms—can be learned correctly in the presence of a dominant pattern.

## References

- Adam Albright and Yoonjung Kang. 2009. Predicting innovative alternations in Korean verb paradigms. *Current issues in unity and diversity of languages: Collection of the papers selected from the CIL*, 18:1–20.

<sup>12</sup>Their time-course analysis tracks this tendency within a single training setting, whereas our comparison spans different suffix conditions. Nevertheless, because all simulations use batch learning and run for the same number of epochs, the results remain broadly comparable.

- Arto Anttila. 2002. Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory*, 20(1):1–42.
- Erwin Chan. 2008. *Structures and distributions in morphology learning*. University of Pennsylvania.
- Andries W Coetzee. 2009. Learning lexical indexation. *Phonology*, 26(1):109–145.
- Sharon Goldwater and Mark Johnson. 2003. Learning of constraint rankings using a maximum entropy model. In *Proceedings of the workshop on variation within Optimality Theory*, pages 111–120.
- Coral Hughto, Andrew Lamont, Brandon Prickett, and Gaja Jarosz. 2019. Learning exceptionality and variation with lexically scaled maxent. In *Proceedings of the Society for Computation in Linguistics (SCIL) 2019*, pages 91–101.
- Sharon Inkelas, C Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. *Derivations and constraints in phonology*, pages 393–418.
- Sharon Inkelas and Cheryl Zoll. 2003. Is grammar dependence real. *Ms, UC Berkeley and MIT*. ROA-587.
- Gaja Jarosz. 2025. Incremental learning of lexically specific morphophonology: an integrative approach. *Linguistics Vanguard*, (0).
- Jongho Jun. 2010. Stem-final obstruent variation in korean. *Journal of East Asian Linguistics*, 19(2):137–179.
- Jordan Kodner. 2022. Computational models of morphological learning. In *Oxford Research Encyclopedia of Linguistics*.
- Ponghyung Lee. 2013. Constraint cloning and lexical listing in korean irregular verbs and adjectives. *Studies in Linguistics*.
- Sechang Lee. 1996. On the irregular verbs in korean. In *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation*, pages 469–478.
- Tal Linzen, Sofya Kasyanenko, and Maria Gouskova. 2013. Lexical and phonological variation in russian prepositions. *Phonology*, 30(3):453–515.
- Joe Pater. 1999. Austronesian nasal substitution and other nc effects. *The prosody-morphology interface*, pages 310–343.
- Joe Pater. 2000. Non-uniformity in english secondary stress: the role of ranked and lexically specific constraints. *Phonology*, 17(2):237–274.
- Joe Pater. 2009. Morpheme-specific phonology: Constraint indexation and inconsistency resolution.
- Sang-Jik Rhee. 1996. The identities of the stem-final consonants of p-and t-irregular verbs in korean. In *Proceedings of the KSPS conference*, pages 476–481. The Korean Society Of Phonetic Sciences And Speech Technology.
- Bruce Tesar and Alan Prince. 2003. Using phonotactics to learn phonological alternations. In *Proceedings from the annual meeting of the Chicago Linguistic Society*, volume 39, pages 241–269. Chicago Linguistic Society.
- Bruce Tesar and Paul Smolensky. 1998. Learnability in optimality theory. *Linguistic inquiry*, 29(2):229–268.
- Yoshimasa Tsuruoka, Jun’ichi Tsujii, and Sophia Ananiadou. 2009. Stochastic gradient descent training for 11-regularized log-linear models with cumulative penalty. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 477–485.

## A Appendix

Category	Stems
p-alternating	neop- ppaep- sep- ssep- gop- bwip- seup- dep- dop- baep- nep- kkeup- moep- kkip- kkop- ssaep- chap- rep- kop- teup- cheop- tep- heup- rwip- jjwip- paep- bop- bup- mip- dwip- choep- sop- swip- pwip- jjeop- keop- soep- ssap- taep- mwip- rup- chwip- bep- peup- ppep- chop- keup- geup- nip- ppip- saep- noep- kkap- jjup- gwip- jjep- meup- ttop- jep- daep- joep- mop- mep- ssoep- ttaep- reop- jjop- ppap-
p-non-alternating	jaep- kkep- jjip- rap- toep-
t-alternating	tat- jat- ket- pot- chit-
t-non-alternating	ppeut- goet- jot- pput- nat- kkut- sit-
s-alternating	ges- kkaes- hoes- koes- nos-
s-non-alternating	poes- meos- jjaes- ches-

Table 12: Toy stems

Condition	Suffix inventory
2-suffix	<i>-ta, -a</i>
4-suffix	<i>-ta, -a, -te, -e</i>
6-suffix	<i>-ta, -a, -te, -e, -tu, -u</i>
8-suffix	<i>-ta, -a, -te, -e, -tu, -u, -to, -o</i>
10-suffix	<i>-ta, -a, -te, -e, -tu, -u, -to, -o, -ti, -i</i>

Table 13: Toy suffixes

## B Appendix

UR	Candidate	% occurrences	*V-OBS-V	MAX	ID-LAT
<i>/tat-a<sub>T</sub>-Alter/</i>	(a) [tata]	0.0	1	0	0
	(b) [tara]	1.0	0	0	1
	(c) [taa]	0.0	0	1	0
<i>/tat-ta<sub>T</sub>-Alter/</i>	(a) [tatta]	1.0	0	0	0
	(b) [tarta]	0.0	0	0	1
	(c) [tata]	0.0	0	1	0
<i>/jot-a<sub>T</sub>-Non-Alter/</i>	(a) [jota]	1.0	1	0	0
	(b) [jora]	0.0	0	0	1
	(c) [joa]	0.0	0	1	0
<i>/jot-ta<sub>T</sub>-Non-Alter/</i>	(a) [jotta]	1.0	0	0	0
	(b) [jorta]	0.0	0	0	1
	(c) [jota]	0.0	0	1	0
<i>/ges-a<sub>S</sub>-Alter/</i>	(a) [gesa]	0.0	1	0	0
	(b) [gera]	0.0	0	0	1
	(c) [gea]	1.0	0	1	0
<i>/ges-ta<sub>S</sub>-Alter/</i>	(a) [gesta]	1.0	0	0	0
	(b) [gerta]	0.0	0	0	1
	(c) [geta]	0.0	0	1	0
<i>/poes-a<sub>S</sub>-Non-Alter/</i>	(a) [poesa]	1.0	1	0	0
	(b) [poera]	0.0	0	0	1
	(c) [poea]	0.0	0	1	0
<i>/poes-ta<sub>S</sub>-Non-Alter/</i>	(a) [poesta]	1.0	0	0	0
	(b) [poerta]	0.0	0	0	1
	(c) [poeta]	0.0	0	1	0

Table 14: Input data tableau for a statistically-balanced condition

UR	Candidate	% occurrences	*V-OBS-V	MAX	ID-C	ID-LAT
<i>/neop-ap-Alter/</i>	(a) [neopa]	0.0	1	0	0	0
	(b) [neowa]	1.0	0	0	1	0
	(c) [neola]	0.0	0	0	0	1
	(d) [neoa]	0.0	0	1	0	0
<i>/neop-tap-Alter/</i>	(a) [neopta]	1.0	1	0	0	0
	(b) [neowta]	0.0	0	0	1	0
	(c) [neolta]	0.0	0	0	0	1
	(d) [neota]	0.0	0	1	0	0
<i>/jaep-ap-Non-Alter/</i>	(a) [jaepa]	1.0	1	0	0	0
	(b) [jaewa]	0.0	0	0	1	0
	(c) [jaela]	0.0	0	0	0	1
	(d) [jaea]	0.0	0	1	0	0
<i>/jaep-tap-Non-Alter/</i>	(a) [jaepta]	1.0	1	0	0	0
	(b) [jaewta]	0.0	0	0	1	0
	(c) [jaelta]	0.0	0	0	0	1
	(d) [jaeta]	0.0	0	1	0	0
<i>/tat-aT-Alter/</i>	(a) [tata]	0.0	1	0	0	0
	(b) [taja]	0.0	0	0	1	0
	(c) [tara]	1.0	0	0	0	1
	(d) [taa]	0.0	0	1	0	0
<i>/tat-taT-Alter/</i>	(a) [tatta]	1.0	0	0	0	0
	(b) [taja]	0.0	0	0	1	0
	(c) [tarta]	0.0	0	0	0	1
	(d) [tata]	0.0	0	1	0	0
<i>/jot-aT-Non-Alter/</i>	(a) [jota]	1.0	1	0	0	0
	(b) [joja]	0.0	0	0	1	0
	(c) [jora]	0.0	0	0	0	1
	(d) [joa]	0.0	0	1	0	0
<i>/jot-taT-Non-Alter/</i>	(a) [jotta]	1.0	0	0	0	0
	(b) [jojta]	0.0	0	0	1	0
	(c) [jorta]	0.0	0	0	0	1
	(d) [jota]	0.0	0	1	0	0
<i>/ges-aS-Alter/</i>	(a) [gesa]	0.0	1	0	0	0
	(b) [geja]	0.0	0	0	1	0
	(c) [gera]	0.0	0	0	0	1
	(d) [gea]	1.0	0	1	0	0
<i>/ges-taS-Alter/</i>	(a) [gesta]	1.0	0	0	0	0
	(b) [gejta]	0.0	0	0	1	0
	(c) [gerta]	0.0	0	0	0	1
	(d) [geta]	0.0	0	1	0	0
<i>/poes-aS-Non-Alter/</i>	(a) [poesa]	1.0	1	0	0	0
	(b) [poeja]	0.0	0	0	1	0
	(c) [poera]	0.0	0	0	0	1
	(d) [poea]	0.0	0	1	0	0
<i>/poes-taS-Non-Alter/</i>	(a) [poesta]	1.0	0	0	0	0
	(b) [poejta]	0.0	0	0	1	0
	(c) [poerta]	0.0	0	0	0	1
	(d) [poeta]	0.0	0	1	0	0

Table 15: Input data tableau for statistically imbalanced condition

## C Appendix

	*V-OBS-V	MAX	ID-LAT
General weights	1.80	3.85	3.85
t-alternating stems	2.76	0.75	0
t-non-alternating stems	0	0	0
s-alternating stems	2.76	0	0.75
s-non-alternating stems	0	0	0
-V suffix	0	0	0
-CV suffix	0	0	0

Table 16: Learned weights (/t, s/ stems; eight suffixes)

	*V-OBS-V	MAX	ID-LAT
General weights	1.72	4.04	4.04
t-alternating stems	3.43	1.14	0
t-non-alternating stems	0	0	0
s-alternating stems	3.43	0	1.14
s-non-alternating stems	0	0	0
-V suffix	0	0	0
-CV suffix	0	0	0

Table 17: Learned weights (/t, s/ stems; ten suffixes)

	*V-OBS-V	MAX	ID-LAT	ID-C
General Weights	6.14	7.76	7.76	3.67
p-alternating stems	0	0	0	0
p-non-alternating stems	0	0	0	3.24
t-alternating stems	1.62	0.0005	0	4.09
t-non-alternating stems	0	0	0	3.24
s-alternating stems	1.62	0	0.0005	4.09
s-non-alternating stems	0	0	0	3.24
-V suffix	0	0	0	0
-CV suffix	0	0	0	0.58

Table 18: Learned Weights (/p, t, s/ stems; eight suffixes)

	*V-OBS-V	MAX	ID-LAT	ID-C
General Weights	7.11	8.98	8.98	4.46
p-alternating stems	0	0	0	0
p-non-alternating stems	0	0	0	3.76
t-alternating stems	2.55	0.69	0	5.2
t-non-alternating stems	0	0	0	3.76
s-alternating stems	2.55	0	0.69	5.2
s-non-alternating stems	0	0	0	3.76
-V suffix	0	0	0	0
-CV suffix	0	0	0	0

Table 19: Learned Weights (/p, t, s/ stems; ten suffixes)

## D Appendix

Two suffixes								
	t-conjugation				s-conjugation			
	t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.523	0.904	0.523	0.904	0.523	0.904	0.523	0.904
P(lateralization)	<b>0.238</b>	0.047	0.238	0.047	0.238	0.047	0.238	0.047
P(deletion)	0.238	0.047	0.238	0.047	<b>0.238</b>	0.047	0.238	0.047

Four suffixes								
	t-conjugation				s-conjugation			
	t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.50	0.93	0.59	0.93	0.50	0.93	0.59	0.93
P(lateralization)	<b>0.25</b>	0.03	0.20	0.03	0.25	0.03	0.20	0.03
P(deletion)	0.25	0.03	0.20	0.03	<b>0.25</b>	0.03	0.20	0.03

Six suffixes								
	t-conjugation				s-conjugation			
	t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.33	0.95	0.73	0.95	0.33	0.95	0.73	0.95
P(lateralization)	<b>0.35</b>	0.024	0.14	0.03	0.31	0.024	0.14	0.02
P(deletion)	0.31	0.021	0.14	0.021	<b>0.35</b>	0.024	0.14	0.02

Eight suffixes								
	t-conjugation				s-conjugation			
	t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.25	0.97	0.79	0.96	0.25	0.97	0.79	0.96
P(lateralization)	<b>0.51</b>	0.10	0.21	0.02	0.24	0.01	0.10	0.02
P(deletion)	0.24	0.01	0.09	0.02	<b>0.51</b>	0.02	0.10	0.02

Ten suffixes								
	t-conjugation				s-conjugation			
	t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.20	0.97	0.84	0.96	0.20	0.987	0.84	0.96
P(lateralization)	<b>0.61</b>	0.02	0.08	0.02	0.19	0.003	0.08	0.02
P(deletion)	0.19	0.01	0.08	0.02	<b>0.61</b>	0.01	0.08	0.02

Table 20: Training-data P(E) under the balanced condition

<b>Two suffixes</b>												
	p-conjugation				t-conjugation				s-conjugation			
	p-alternating		p-non-alternating		t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.181	0.968	0.181	0.968	0.181	0.968	0.181	0.968	0.181	0.968	0.181	0.968
P(glide)	<b>0.713</b>	0.011	0.713	0.011	0.713	0.011	0.713	0.011	0.713	0.011	0.713	0.011
P(lateralization)	0.053	0.011	0.053	0.011	<b>0.053</b>	0.011	0.053	0.011	0.053	0.011	0.053	0.011
P(deletion)	0.053	0.011	0.053	0.011	0.053	0.011	0.053	0.011	<b>0.053</b>	0.011	0.053	0.011

<b>Four suffixes</b>												
	p-conjugation				t-conjugation				s-conjugation			
	p-alternating		p-non-alternating		t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.123	0.976	0.313	0.986	0.313	0.986	0.313	0.986	0.313	0.986	0.313	0.986
P(glide)	<b>0.803</b>	0.013	0.497	0.003	0.497	0.003	0.497	0.003	0.497	0.003	0.497	0.003
P(lateralization)	0.037	0.005	0.095	0.005	<b>0.095</b>	0.005	0.095	0.005	0.095	0.005	0.095	0.005
P(deletion)	0.037	0.005	0.095	0.005	0.095	0.005	0.095	0.005	<b>0.095</b>	0.005	0.095	0.005

<b>Six suffixes</b>												
	p-conjugation				t-conjugation				s-conjugation			
	p-alternating		p-non-alternating		t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.089	0.982	0.434	0.995	0.334	0.995	0.434	0.995	0.334	0.995	0.434	0.995
P(glide)	<b>0.863</b>	0.014	0.332	0.001	0.333	0.001	0.332	0.001	0.333	0.001	0.332	0.001
P(lateralization)	0.024	0.002	0.117	0.002	<b>0.166</b>	0.002	0.117	0.002	0.166	0.002	0.117	0.002
P(deletion)	0.024	0.002	0.117	0.002	0.166	0.002	0.117	0.002	<b>0.166</b>	0.002	0.117	0.002

<b>Eight suffixes</b>												
	p-conjugation				t-conjugation				s-conjugation			
	p-alternating		p-non-alternating		t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.076	0.985	0.538	0.998	0.25	0.998	0.538	0.998	0.25	0.998	0.537	0.998
P(glide)	<b>0.894</b>	0.014	0.249	0.0005	0.25	0.0002	0.249	0.0005	0.25	0.0002	0.249	0.0005
P(lateralization)	0.015	0.0004	0.106	0.0004	<b>0.249</b>	0.0004	0.106	0.0004	0.249	0.0004	0.106	0.0004
P(deletion)	0.015	0.0004	0.106	0.0004	0.249	0.0004	0.106	0.0004	<b>0.249</b>	0.0004	0.106	0.0004

<b>Ten suffixes</b>												
	p-conjugation				t-conjugation				s-conjugation			
	p-alternating		p-non-alternating		t-alternating		t-non-alternating		s-alternating		s-non-alternating	
	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV	-V	-CV
P(faithful)	0.064	0.988	0.61	0.999	0.2	0.999	0.61	0.999	0.2	0.999	0.611	0.999
P(glide)	<b>0.914</b>	0.011	0.20	0.0002	0.2	0	0.2	0	0.2	0	0.2	0
P(lateralization)	0.010	0.0001	0.094	0.0001	<b>0.398</b>	0	0.094	0	0.2	0	0.094	0
P(deletion)	0.010	0.0001	0.094	0.0001	0.2	0	0.09	0	<b>0.398</b>	0.0001	0.09	0

Table 21: Training-data P(E) under the imbalanced condition

## E Appendix

p-alternating 45 (45%)
p-non-alternating 17 (17%)
t-alternating 8 (8%)
t-non-alternating 14 (14%)
s-alternating 8 (8%)
s-non-alternating 8 (8%)

Table 22: Distribution of /p/, /t/, and /s/-final stems in the Google Treebank Corpus

Category	Stems
p-alternating	gakkap- gomap- gop- gup- geurip- deop- dop- dukkeop- tteugeop- maep- mugeop- museop- budeureop- saerop- swip- sikkeureop- akkap- areumdap- aswip- antakkap- eoryeop- jeulgeop- chup- himgyeop- kkadarop- deoreop- maekkeureop- bangap- bukkeureop- sonswip- singgeop- useup- jeonggyeop- gabyeop- goerop- gwiyeop- nollap- nup- duryeop- duteop- maeseop- ansseureop- eodup- oerop- chagap-
p-non-alternating	dwijip- butjap- ppob- ssip- ip- jap- job- barojap- hwijip- kkojip- kkob- sarojap- sonjap- jip- jjip- charyeip- himip-
t-alternating	geot- deut- mut- sit- kkaedat- ilkkeut- git- but-
t-non-alternating	dat- mit- bat- ssot- eot- gut- mulryeobat- beorimbat- injeongbat- geot- got- tteut- mut- ieobat-
s-alternating	nas- bus- is- jis- maedeupjis- nongsajis- dwis- yeonis-
s-non-alternating	ssis- us- johas- chisos- beos- byas- ppaeas- is-

Table 23: /p/, /t/, and /s/-final stems in the Google Treebank Corpus (romanized)<sup>a</sup>

<sup>a</sup> Romanization follows [https://www.korean.go.kr/front\\_eng/roman/roman\\_01.do](https://www.korean.go.kr/front_eng/roman/roman_01.do).

Context	Suffix inventory
-CV	-geona -ge -get -go -goyo -guyo -gi -gillae -na - nayo -neyo -neureago -neun -neunda -neundago -neunde -da -daga -dago -daneun -deogunyo - deon -dorok -seumnida -ja -janeun -jamaja -ji -jiman
-V	-n -nde -l -m -myeo -myeon -myeonse -a -ado -aseo -aya -ayo -at -eo -eodo -eoseo -eoya - eoyo -eot -eotseot -euna -euni -eureo -eureogo -eureoneun -eumyeo -eumyeon -eumyeonse - euseyo -eusi -eun -eul -eum

Table 24: Suffix inventories in the Google Treebank Corpus (romanized)<sup>a</sup>

<sup>a</sup> Romanization follows [https://www.korean.go.kr/front\\_eng/roman/roman\\_01.do](https://www.korean.go.kr/front_eng/roman/roman_01.do).