General-over-specific Markedness Bias as a Balancing Force in GLA-style Learning

Kaili Vesik
University of British Columbia

1 Introduction

Learning can be tricky, especially when attempting to model the acquisition of one language in the context of an entire typology. Constraints with overlapping violation profiles can cause challenges (Credit Problem; Dresher, 1999) for a Gradual Learning Algorithm (GLA; Boersma & Hayes, 2001) type learner. These are exacerbated when marrying stringency scales (de Lacy, 2002) with no-disagreement constraints (Pulleyblank, 2002), producing a complex, overlapping set. In this paper I propose a novel general-over-specific markedness ($\mathrm{M_{gen}} \gg \mathrm{M_{spec}}$) bias, not previously thought to be necessary for this kind of learner, to address a restrictiveness problem that is made evident in this learning context. I present the simulated acquisition of North Estonian vowel patterns (situated in the Finnic typology) as a case where introducing a $\mathrm{M_{gen}} \gg \mathrm{M_{spec}}$ bias as counterpart to specific-over-general faithfulness helps to offset such adverse effects.[1] Such a bias helps to ensure maximal restrictiveness in the final grammar acquired by the learner.

I begin by introducing the necessary background for modeling acquisition of Finnic languages, with a particular focus on North Estonian: Section 2.1 describes the typology of Finnic vowel patterns and Section 2.2 proposes a constraint set to account for those patterns. Next, in Section 3 I present the basic setup for a gradual error-driven learner simulating acquisition of these languages, followed by specifications and results for a first pass at an instantiation of such a learner. Section 4 presents a second instantiation of a Finnic learner, whose results demonstrate the advantage of including a general-over-specific markedness bias. Finally, Section 5 concludes.

2 Background

2.1 *Typology* The Finnic languages (Uralic family) are spoken in Scandinavia and northern Europe (Bakró-Nagy et al., 2022). Their vowel inventories are drawn from the following mutually exclusive sets of phonologically front vowels /i, e, æ, y, ø/ and back vowels /ɯ, ɤ, ɑ, u, o/; see Figure 1.

The language descriptions in this section have been gathered and synthesized from sources that cover a broad spectrum in terms of both time span and transcription conventions. In order to present the facts using consistent notation, I have abstracted the vowel descriptions to these ten IPA (International Phonetic Alphabet) symbols. Though they are not always the same symbols used by the authors of the original literature, they achieve the same ends of partitioning the vowel space

* **My thanks to Anne-Michelle Tessier, Gunnar Ólafur Hansson, Kathleen Currie Hall, participants at AMP 2023, the UBC Phonology Discussion Group, and the UBC Child Phonology Lab. This research is funded in part by a SSHRC CGS-D.**

[1] This poster was presented fifteen months ago, and readers may notice that the approach I proposed at that time was rather different than this one. The solution I offer herein is, I believe, a vast improvement on my earlier hypotheses.

Proceedings of *AMP 2023/2024*

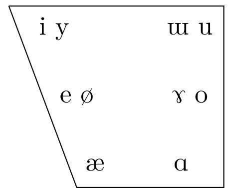

Figure 1: Finnic vowel space.

both by height (high, mid, and low) and advancement (front and back). Note also that for the sake of simplicity, I include transcriptions only for the *vowels* (omitting the consonants), and only for vowel *quality* (omitting any references to quantity).

The Finnic languages summarized in Table 1 define a typological landscape of vowel monophthong distribution patterns that includes various degrees of (a) positional restrictions and (b) progressive back-front vowel harmony (VH). The languages that exhibit harmony may include harmonic alternations, neutral transparent vowels, and/or neutral opaque vowels. In this paper I focus on North Estonian, which has an inventory of nine vowels, restricted distribution in non-initial syllables, and no vowel harmony.

Languages	**Inventory**	**Restrictions outside of σ_1**	**Vowel harmony**			
			Back	**Front**	**Transp**	**Opaque**
N Estonian	i,e,æ,y,ø,ɤ,ɑ,u,o	* æ,y,ø,ɤ	-	-	-	-
Livonian	i,e,æ,ɯ,ɤ,ɑ,u,o	* æ,ɯ,ɤ,o	-	-	-	-
Finnish Karelian Ingrian	i,e,æ,y,ø,ɑ,u,o	-	ɑ,u,o	æ,y,ø	i,e	-
Votic Kihnu Est	i,e,æ,y,ø,ɤ,ɑ,u,o	-	ɤ,ɑ,u,o	e,æ,y,ø	i,(e)	-
N Seto	i,e,æ,y,ø,ɯ,ɤ,ɑ,u,o	* ɯ	ɯ,ɤ,ɑ,u,o	e,æ,y,ø	i	-
S Seto	i,e,æ,y,ø,ɯ,ɤ,ɑ,u,o	* ø,ɯ	ɯ,ɤ,ɑ,u,(o)	e,æ,y,ø	i,(e)	o
Veps	i,e,æ,y,ø,ɑ,u,o	(* æ,y,ø)	ɑ,u,o	æ,y,ø	i,e	-

Table 1: Selected BF vowel patterns, synthesized from Ariste (1968); Campbell et al. (2013); Fejes et al. (2024); Grünthal (2015, 2022); Karlsson (2018); Kiparsky & Pajusalu (2003); Laakso (2022); Léonard (1993); Markus & Rozhanskiy (2022a,b); Metslang (2022); Nikolaev (2018, 2019); Pajusalu (2022); Sang (2009); Sarhimaa (2022); Suomi et al. (2008).

2.2 *OT analysis* In order to account for the patterns summarized in Section 2.1, I use a framework that combines elements from Pulleyblank (2002) and de Lacy (2002). These components are used with the intent to (a) ensure that vowel harmony is feasible where applicable, and (b) facilitate the scales of both markedness (in terms of non-initial vowel restrictions) and harmony participation (in terms of neutral vowels).

Pulleyblank (2002) proposes a *no-disagreement* approach to harmony which, rather than focusing on agreement or alignment/spreading of harmonic features as the driver for harmony, focuses on avoiding disagreement via banned sequences. For example, in a front-back harmony system this approach might include constraints such as *[+back][-back] and *[-back][+back] rather than one like Agree(Back). Pulleyblank also proposes a *focus* for each constraint, marked by an underlined value; this focus is the feature value for which violations are counted. That is, even though there are many unique [b]…[f] pairs in candidate [b..f..f..f], it only incurs one violation of *[+back]…[-back] because there is only one back vowel to which to assign violations.

De Lacy (2002) proposes an approach to markedness and scales/hierarchies following Prince's

(1997 et seq., as cited by de Lacy) *stringency relations*. He uses hierarchical markedness scales as well as *scale-referring constraints*. These hierarchical structures achieve similar ends as fixed rankings do in other systems, while allowing for free permutation of constraints.

The use of stringency scales is advantageous in accounting for the varying degrees of both positional restrictions and vowel harmony in Finnic languages. For positional restrictions, since some back vowels (e.g. /ɯ, ɤ/) and front vowels (e.g. /æ, y, ø/) are consistently more restricted than others, we can build scales to reflect these degrees of markedness. For vowel harmony, since some front vowels (e.g. /i, e/) are less likely to participate in harmony than others, scales can reflect these degrees of participation as well, working in concert with no-disagreement constraints to allow for some types of disharmony (e.g. [ɑ..i..u] as a back-harmonic sequence containing a front transparent vowel) but ban others (e.g. *[ɑ..ø..u] as a back-harmonic sequence containing a front harmonic vowel).

Before presenting constraints, I define the stringency scales that form the foundation of the system; see Table 2. The vowel sets defined in the table are referred to by both the context-free and the no-disagreement constraints (3) through (8) below.

Vowel class	Markedness scale	Scale-referring sets
Front	(a) ø > æ,y > e > i[2]	i. F_1 = {ø} ii. F_2 = to be determined iii. F_3 = {ø,æ,y} iv. F_4 = {ø,æ,y,e} v. F_5 = {ø,æ,y,e,i}
Back	(b) ɯ > ɤ > o > ɑ,u	i. B_1 = {ɯ} ii. B_2 = {ɯ,ɤ} iii. B_3 = {ɯ,ɤ,o} iv. B_4 = to be determined v. B_5 = {ɯ,ɤ,o,ɑ,u}

Table 2: Vowel sets as stringency scales.

2.2.1 *Constraints* My approach is concerned only with markedness repairs involving changes to the feature [back]. As such, I begin the introduction of my constraint set by proposing two faithfulness constraints:

(1) Iᴅ(Bk): Assign a violation mark for each segment in the output whose input correspondent has a different value for [back].

(2) Iᴅ-σ_1(Bk): Assign a violation mark for a first-syllable vowel in the output whose input correspondent has a different value for [back].

I assume Iᴅ-σ_1(Bk) $\gg$ Iᴅ(Bk). This ensures that the first syllable's vowel (a) is eligible for full contrast in a language with positional restrictions, and/or (b) can drive harmony in a language with vowel harmony.

The scale-referring context-free markedness constraints take one of the two forms below (with $m \in \{1, 3, 4, 5\}$ and $n \in \{1, 2, 3, 5\}$):

(3) *F_m: Assign a violation mark for each segment in the output that is in F_m.

(4) *B_n: Assign a violation mark for each segment in the output that is in B_n.

[2] De Lacy (2002) argues that scales must have individual elements strictly ordered, in order to avoid potential pathological effects stemming from /æ, y/ being grouped as if in a natural class. However, I do not yet have enough evidence to determine whether the ordering here is æ > y or y > æ, so for the time being I retain this ambiguous subsequence. The same issue arises with the back vowels as well.

The scale-referring harmony (no-disagreement) markedness constraints all take one of the four forms below (with $m, q \in \{1, 3, 4, 5\}$ and $n, p \in \{1, 2, 3, 5\}$):[3]

(5) $^*\mathrm{F}_m\underline{\mathrm{B}}_n$: Assign a violation for each instance of a back vowel from set B_n, immediately preceded by a front vowel from set F_m.

(6) $^*\mathrm{F}_m...\underline{\mathrm{B}}_n$: Assign a violation for each instance of a back vowel from set B_n, preceded at any distance by at least one front vowel from set F_m.

(7) $^*\underline{\mathrm{B}}_p\mathrm{F}_q$: Assign a violation for each instance of a back vowel from set B_p, immediately followed by a front vowel from set F_q.

(8) $^*\underline{\mathrm{B}}_p...\mathrm{F}_q$: Assign a violation for each instance of a back vowel from set B_p, followed at any distance by at least one front vowel from set F_q.

I have not enumerated all of the possible combinations of the markedness constraints, as there would be eight context-free constraints and $4 \times 4 = 16$ of each form in (5) through (8) above, for a total of 64 no-disagreement constraints. Theoretically this could mean tableaux of enormous size, but owing to the fact that many of these constraints are inactive (see Section 2.2.2), we can generally focus on just the handful that are relevant to any given grammar.

2.2.2 *Application to North Estonian* North Estonian has two main vowel phenomena that must be accounted for: inventory gaps and positional restrictions. The only Finnic vowel that is completely missing from North Estonian is /ɯ/; that is, the set B_1. Therefore it must be the case that $^*\mathrm{B}_1$ outranks both faithfulness constraints:

(9) $^*\mathrm{B}_1 \gg$ Id-σ_1(Bk)$\gg$ Id(Bk)

In terms of positional restrictions, vowels /ø, æ, y, ɯ, ɤ/ (that is, those in sets F_3 and B_2) appear in initial syllables but not in non-initial syllables. Thus $^*\mathrm{F}_3$ and $^*\mathrm{B}_2$ must be sandwiched between the specific and general faithfulness constraints, to permit these vowels to surface in initial syllables but not later:

(10) Id-σ_1(Bk) $\gg$ $^*\mathrm{F}_3$, $^*\mathrm{B}_2$ $\gg$ Id(Bk)

Tableau (11) shows that both:

- Id-σ_1(Bk) must outrank $^*\mathrm{B}_2$ in order to ensure that the first syllable's vowel is preserved, even if it falls into this marked set, and
- $^*\mathrm{B}_2$ must outrank Id(Bk) in order to ensure that marked vowels do not surface past the first syllable of the word, even at the expense of faithfulness to the underlying vowel.

Parallel arguments can be made for $^*\mathrm{F}_3$.

(11)

/ɤ..ɤ/	Id-σ_1(Bk)	$^*\mathrm{B}_2$	Id(Bk)
a. ɤ..ɤ		**!	
☞ b. ɤ..e		*	*
c. e..ɤ	*!	*	*
d. e..e	*!		**

Overall, then, the rankings for North Estonian are as depicted in the Hasse diagram in Figure 2. The relative rankings following the leftmost path are the ones justified above. The relationships between the remaining constraints can be summarized as follows:

- $^*\mathrm{F}_1$ must rank below Id-σ_1(Bk) so that vowels from this set are not banned altogether.

[3] Note that for the harmony constraints, I assume access to a vowel tier (Goldsmith, 1976; Clements, 1976).

- $*F_{4,5}$ and $*B_{3,5}$ must rank below $*F_3$ and $*B_2$ so that only vowels from the smaller sets are banned in non-initial syllables.
- $*\underline{B}_iF_{4,5}$, $*\underline{B}_i{\dots}F_{4,5}$, $*F_j\underline{B}_{3,5}$, and $*F_j{\dots}\underline{B}_{3,5}$ must rank below $*F_3$ and $*B_2$ so that only vowels from the smaller sets are banned in non-initial syllables.
- $*\underline{B}_iF_{1,3}$, $*\underline{B}_i{\dots}F_{1,3}$, $*F_j\underline{B}_{1,2}$, and $*F_j{\dots}\underline{B}_{1,2}$ are unranked with respect to the rest of the constraints. These no-disagreement constraints refer to the restricted vowel sets in their second elements and are therefore never violated by the North Estonian data.

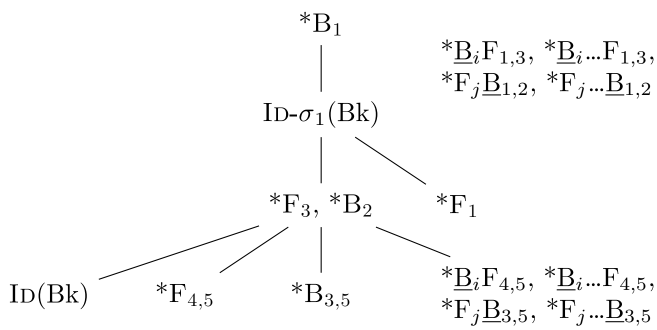

Figure 2: Overall North Estonian rankings

These rankings represent the target grammar for the learning simulations described in Sections 3 and 4.

3 Learning without $\mathbf{M_{gen} \gg M_{spec}}$

Simulating acquisition of a grammar with an algorithmic learner involves many potential variables, parameters, and biases. In this section I introduce the general algorithm used as a base for learning Finnic vowel patterns, and present results from a simulation with North Estonian inputs.

3.1 *A basic GLA-type learner*

3.1.1 *Description* The GLA (Boersma & Hayes, 2001) forms the basis for the gradual, error-driven learners that follow. In particular:

- The learner tracks and adjusts numeric values assigned to constraints. These *ranking values* (θ_C for each constraint C) determine the constraints' relative ordinal rankings at evaluation time so that inputs can be evaluated with a classic OT grammar.
- For each learning trial, the learner ranks all constraints in order of their current ranking values (modulo a small amount of noise) to produce a current hypothesized grammar. This grammar is used to determine the optimal output for the given input.
- If the optimal candidate is not the intended winner (i.e. the optimal candidate is an intended loser), the current grammar is updated. Winner-preferring constraints are promoted and loser-preferring constraints are demoted.

This type of learner assumes that given enough positive data, making small changes in response to one input at a time will eventually get the learner to the target grammar. However, this optimistic assumption also has the potential to send the learner trundling down the wrong path. If the learner happens to make an inaccurate assumption (crucially, one that puts the learner on a trajectory toward a less restrictive grammar) about the phonotactics motivating a particular type of pattern, there is no amount of positive data that can save the learner from acquiring a grammar that overgenerates. This kind of hazard is made more concrete in Section 3.2.3.

3.1.2 *Data and parameter settings* In terms of learning data, in these simulations I use positive data only, consisting of all vowel sequences of lengths 2 and 3 that are attested in native non-compound North Estonian words. Inputs are assumed to be identity-mapped; that is, any surface form is assumed to have the same underlying form. The simulations are run using my own implementation of the algorithm in Python (Van Rossum & Drake, 2009). However, some of the default settings discussed below are drawn from OTSoft (Hayes et al., 2013).

The number of learning trials is fixed at $20,000$ for each simulation. Both simulations described herein converged well before iterating through this many trials, providing a long enough timeline to ensure that even the odd later error (caused by a particularly noisy evaluation) did not affect the overall ranking. The organization of learning trials into stages, evaluation noise, and the plasticity function are summarized in Table 3.

Parameter	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**
Number of learning trials	5000	5000	5000	5000
Evaluation noise	2	2	2	2
Plasticity	2	0.2	0.02	0.002

Table 3: Numerical parameters for GLA-type learning.

The bare bones of the learning algorithm as described above lay the foundation for additional potential parameters or biases to be included in any instantiation of a learner. Section 3.2 describes Learner A (the first of two such learners to be discussed in this paper), including implementation, results, and analysis of shortcomings. Section 4 investigates Learner B, involving an additional bias, which produces much improved results as compared to Learner A.

3.2 *Learner A*

3.2.1 *Description and implementation* Learner A is defined with the following biases in addition to the basic setup described above in Section 3.1.2.

Low-faithfulness bias: The bias toward low initial faithfulness is widely used in the learning literature, as it helps to ensure that the acquired grammar is as restrictive as possible; that is, it mitigates the Subset Problem (Angluin, 1980; Baker, 1979). Readers can find more detailed discussion in, e.g., Gnanadesikan (1995); Smolensky (1996); Hayes (2004); Prince & Tesar (2004); Jesney & Tessier (2011).

I use a default implementation in which the initial ranking value of faithfulness constraints is 0, and that of markedness constraints is 100. There is another bias discussed in Section 4.1 that sets initial markedness values to be different from the default; however, these will continue to preserve the overarching low-faithfulness bias.

Specific-over-general faithfulness bias: The constraint set that I use for this project includes only two faithfulness constraints, ID(Bk) and ID-σ_1(Bk), the first applying more broadly and the second in a narrower context. When two such versions of a faithfulness constraint exist, it is possible to construct a grammar in which marked elements in underlying forms surface only in privileged contexts. For example, recall that the ranking ID-σ_1(Bk) $\gg$ *F_3 $\gg$ ID(Bk) bans vowels in set F_3 in general, but permits them in initial syllables.

A specific-over-general faithfulness bias ($F_{spec} \gg F_{gen}$) is a strategy that can help find the most restrictive grammar that accounts for the input data, avoiding a superset (overgenerating) grammar (Hayes, 2004; Tessier, 2007). Such a bias ensures that the specific version of the constraint always has a better opportunity to claim credit for a particular output form than the general one does, corresponding to a more restrictive grammar overall.

The $F_{spec} \gg F_{gen}$ bias between any specific-general pair of faithfulness constraints can be implemented by means of an *a priori* bias that ensures $\theta_{F_{spec}} - \theta_{F_{gen}} \geq d$, for some distance d. Practically, the learner adjusts the initial ranking values such that any two constraints in this type of relationship are at least d apart, and then does the same after each learning update. If the two constraints have a difference of less than d, then it is always the case that the specific one has its value increased rather than the general one having its value decreased. I use the OTSoft (Hayes et al., 2013) default value of $d = 20$, which is described as "very close probabilistically to being an obligatory ranking" (Hayes, 2013: 24).

Tempered promotion rate: GLA-type learners make adjustments to the ranking values after each error made by the current hypothesized grammar. While all variations on this theme agree that constraint demotion is necessary to the learning process, arguments have been made both for (e.g., Boersma, 1997, 1998; Magri, 2012) and against (e.g., Tesar & Smolensky, 1998) the idea of permitting constraint promotion as well. I subscribe to Magri's (2012) claim that some promotion must be required in order to allow for re-ranking of faithfulness constraints which, in a learning environment that assumes faithful underlying forms for licit inputs, never prefer losers.

The amount that winner-preferring constraints get promoted can be calculated as a fraction of the current plasticity (Magri, 2012), with the fraction determined as a function of the number of winner-preferring and/or loser-preferring constraints at that update. For example,

(12) $$\text{promotion amount} = \text{promotion rate} \times \text{plasticity}$$

At the low end, a promotion rate of 0 means that initially-low faithfulness constraints would remain stuck at their starting values; they need some way to rise to allow for adjustments to rankings as new learning inputs are encountered. At the high end, a promotion rate of 1 (or more) means that every winner-preferring constraint is given full credit for preference of the winner. However, we should consider avoiding full-fledged promotion of constraints in the case of an update where two or more constraints prefer the intended winner, in order to avoid overpromoting when it is not clear which of those constraints should be credited with preference of the winner (Credit Problem; Dresher, 1999).

With this in mind, at each learning update, the ranking values of the loser-preferrers are decreased by the plasticity amount, and those of the winner-preferrers are increased by the promotion rate as a fraction of the plasticity. I use Magri & Kager's (2015) promotion rate of $\frac{1}{1+W}$, where W = the number of winner-preferring constraints involved in the update. Thus:

(13) $$\text{promotion amount} = \frac{1}{1+W} \times \text{plasticity}$$

3.2.2 *Results* To demonstrate the effect of this initial set of biases, I simulate acquisition of North Estonian using Learner A, defined with the settings described in Sections 3.1.2 and 3.2.1 above. The final grammar produced by this learner is reasonably good, with a success rate of 0.7865 on test evaluations.[4] However, this is not nearly at the level that a learner should be if it is to convincingly mirror human acquisition. We can get a better sense of what is amiss by analyzing the final ranking values produced by the learner and inspecting the Elementary Ranking Conditions (ERCs; Prince, 2002) associated with relevant updates. These final ranking values, for a selection of crucial constraints, are shown in Table 4. ERCs and discussion follow.

Constraint	**Final ranking value**
$*\mathrm{B}_1$	100.000
$*\mathrm{F}_5\underline{\mathrm{B}}_2$	100.000
$*\mathrm{F}_5...\underline{\mathrm{B}}_2$	100.000
$*\underline{\mathrm{B}}_5\mathrm{F}_3$	100.000
$*\underline{\mathrm{B}}_5...\mathrm{F}_3$	100.000
Id-σ_1(Bk)	83.021
$*\mathrm{B}_2$	63.267
Id(Bk)	63.021
$*\mathrm{F}_3$	23.457

Table 4: Final ranking values of select constraints after simulation with Learner A.

[4] See Vesik (in prep) for a detailed description of testing methods.

3.2.3 *Discussion* Referring to Table 4, the highly-ranked $*B_1$ successfully accounts for an inventory gap with respect to /ɯ/. However, the positional restrictions are more problematic. The final grammar shows that this learner, due in part to only having access to positive evidence, has "misunderstood" North Estonian as being a vowel harmony language. $*B_2$'s final ranking value is just barely greater than Id(Bk)'s, which means that the two have almost a 50% likelihood of switching places due to evaluation noise. $*F_3$, on the other hand, has finished well below Id(Bk) and is therefore inactive. The restricted F_3 vowels in non-initial syllables are accounted for by no-disagreement constraints $*\underline{B}_5F_3$ and $*\underline{B}_5…F_3$ rather than the context-free constraint $*F_3$, since those no-disagreement constraints are never violated while $*F_3$ is violated often by vowels in the first syllable. The ERC matrix in Table 5 shows that when the input happens to contain two relatively unmarked back vowels (/u..ɑ/), both $*\underline{B}_5…F_3$ and $*F_3$ are winner-preferring and therefore promoted, but when the input sequence contains a marked front vowel (/y..ɑ/), $*F_3$ is demoted while $*\underline{B}_5…F_3$ does not move.

input	candidates	Id-σ_1(Bk)	Id(Bk)	$\boldsymbol{*F_3}$	$\boldsymbol{*\underline{B}_5…F_3}$	$\boldsymbol{*B_5}$
/u..ɑ/ →	u..ɑ∼u..æ	e	L	W	W	L
/y..ɑ/ →	y..ɑ∼u..ɑ	W	W	L	e	W

Table 5: ERC matrix comparing inputs with marked vs unmarked vowels in the initial syllable.

Interpreting North Estonian as a vowel harmony language rather than one with broader positional restrictions means that vowels in set F_3 (and sometimes B_2) are banned only when following a vowel of the opposite backness. Hence sequences such as [y..ø] are deemed acceptable even though such sequences are not attested in North Estonian. Tableaux (14) and (15) show how vowels in set B_2 might be appropriately restricted in non-initial syllables but those in set F_3 are definitely not.

(14)

/ɑ..ɤ/	$*F_5\underline{B}_2$	$*F_5…\underline{B}_2$	Id-σ_1(Bk)	$*B_2$	Id(Bk)
a. ɑ..ɤ				*!	
☞ b. ɑ..e					*
c. æ..ɤ	*!	*	*	*	*
d. æ..e			*!		**

(15)

/y..ø/	$*\underline{B}_5F_3$	$*\underline{B}_5…F_3$	Id-σ_1(Bk)	Id(Bk)	$*F_3$
☹ ☞ a. y..ø					*
b. y..o				*!	*
c. u..ø	*!	*	*	*	*
d. u..o			*!		**

Learner A acquires a North Estonian grammar that accounts for the patterns attested in the positive-evidence learning inputs via no-disagreement constraints rather than context-free markedness constraints, since these particular no-disagreement constraints are never violated while the context-free constraints are adjusted downward as they are repeatedly violated by the learning data. The problem with such a result is that the no-disagreement constraints are too specific– they account perfectly well for the input data but fail to generate correct outputs when given ungrammatical test data. The vowels in non-initial syllables need to be restricted not only when they are disharmonic with the vowels in initial syllables, but also even when they harmonize. This is in essence a restrictiveness problem: in the same way that prioritizing specific faithfulness constraints over general ones allows for a more restrictive grammar, it is clear here that general markedness constraints need to be prioritized over specific ones in order to ensure better restrictiveness in terms of markedness as well.

There is a wide range of potential strategies for prioritizing generality in markedness constraints, which are detailed in Vesik (in prep). Section 4 presents a novel implementation of one such strategy that is particularly effective in this case, along with simulation results.

4 Learning with $\mathrm{M_{gen}} \gg \mathrm{M_{spec}}$

4.1 *Learner B*

4.1.1 *Description* The constraint set defined in Section 2.2.1 does not have any explicitly defined pairs of specific vs general markedness constraints like it does faithfulness constraints. However, it is nevertheless possible to determine the relative generality of the markedness constraints. Using a general-over-specific markedness bias works toward the same goal as the specific-over-general faithfulness bias: learning a grammar that is as restrictive as possible.

Learner A has difficulty learning a correct ranking for North Estonian because the vowel harmony (no-disagreement) constraints do a better job of explaining the limited vowels in non-initial syllables in the learning data than the context-free segmental constraints do. Although the context-free constraints are sometimes violated (specifically, by vowels in the first syllable), in fact they are much more general and can better deal with ungrammatical inputs than the no-disagreement constraints can. The more general constraints make the grammar more restrictive, which is useful for avoiding overgeneration when encountering ungrammatical inputs.

The rationale for a general-over-specific markedness bias is to give the most general markedness constraints an opportunity to get credit for the phonotactics of the target grammar, in order to ensure maximal restrictiveness. The preference for more general markedness constraints is not persistent; rather, it is implemented as an initial articulated hierarchy of markedness constraint values, calculated as a function of each constraint's rate of application in a sample set of inputs (cf. Albright & Hayes, 2006: 11) and can be freely reversed by learning data.

4.1.2 *Implementation* While there are several different dimensions on which generality of the scale-referring markedness constraints can be measured theoretically, informed by set theory (Vesik, in prep), here I calculate the initial ranking values of markedness constraints via a numerical method based on the application rate of each constraint within the inputs seen by the learner. This distribution function was selected for its minimal requirement of *a priori* knowledge or calculation on the part of the learner, given that the initial distribution of markedness constraints can be determined based strictly on observation rather than prior analysis of set-theoretic relationships between classes of markedness constraints.

Recall that in Section 3.1 I present the idea of the learning process taking place over four stages, each with a declining plasticity and consisting of 5000 learning trials. For the implementation of this particular bias, I prepend a pre-learning "observation" stage, with plasticity = 0.

During the observation stage, the learner is fed randomly-sampled inputs just as it is during the learning stages. However, rather than using the current hypothesized grammar to compare the optimal output to the intended winner, the learner simply observes how many times each markedness constraint is violated by the candidate corresponding to the heard input, and adds it to the tally for that constraint. Once the learner has heard all of the inputs in this stage, the violations tally for each constraint is divided by the number inputs heard, producing an average generality measure (application rate) for each markedness constraint.

Before the first learning stage, the generality is used to calculate the initial ranking value for each markedness constraint, modifying it from the default value of 100. This is done using Equation 16.

(16)
$$\begin{aligned}
\theta_{\mathrm{M_{init}}} &= 100(b + m \cdot g_{\mathrm{M}}), \text{ where} \\
b &= y\text{-intercept coefficient} \\
&\quad (\text{determines } \theta_{\mathrm{M_{init}}} \text{ for a constraint with } g_{\mathrm{M}} = 0.0) \\
m &= \text{slope coefficient} \\
&\quad (\text{determines } \theta_{\mathrm{M_{init}}} \text{ for a constraint with } g_{\mathrm{M}} = 1.0) \\
g_{\mathrm{M}} &= \text{generality for constraint M}
\end{aligned}$$

4.1.3 *Results* To demonstrate the cumulative effect of the initial set of biases along with the $\mathrm{M_{gen}} \gg \mathrm{M_{spec}}$ bias, I simulate acquisition of North Estonian using Learner B, defined with the settings described in Sections 3.1.2, 3.2.1, and 4.1.1 above. The initial ranking values for a subset of the constraints are shown in Table 6a. These constraints are selected in order to demonstrate the wide range of initial values for markedness constraints based on their relative generality; note that they are not exaclty the same ones that are shown in the final rankings table. The final ranking values, on the other hand, for a selection of crucial constraints is shown in Table 6b. The final grammar produced by this learner shows vast improvement over the one produced by Learner A, with a success rate of 0.9837 on test evaluations. It meets all the requirements for a target North Estonian grammar (recall Figure 2); however, there is one subtle outstanding issue with this grammar, which is discussed below.

Constraint	**Initial ranking value**
$^*\mathrm{B}_5$	254.74
$^*\mathrm{F}_5$	228.66
$^*\mathrm{F}_4$	179.44
$^*\mathrm{B}_3$	158.26
$^*\mathrm{F}_3$	132.26
$^*\mathrm{B}_2$	110.68
$^*\mathrm{B}_1$	100.00
$^*\mathrm{F}_5\underline{\mathrm{B}}_2$	100.000
$^*\underline{\mathrm{B}}_5\mathrm{F}_3$	100.000

(a) Initial ranking values of selected constraints prior to simulation with Learner B.

Constraint	**Final ranking value**
$^*\mathrm{B}_1$	110.667
$^*\mathrm{F}_5\underline{\mathrm{B}}_2$	100.000
$^*\mathrm{F}_5\ldots\underline{\mathrm{B}}_2$	100.000
$^*\underline{\mathrm{B}}_5\mathrm{F}_3$	100.000
$^*\underline{\mathrm{B}}_5\ldots\mathrm{F}_3$	100.000
ID-σ_1(Bk)	96.941
$^*\mathrm{B}_2$	88.243
$^*\mathrm{F}_3$	80.563
ID(Bk)	76.941

(b) Final ranking values of selected constraints after simulation with Learner B.

Table 6: Comparison of initial vs final Learner B ranking values for selected constraints.

4.1.4 *Discussion* Referring to Table 6b, both the inventory gap (recall (9): $^*\mathrm{B}_1 \gg$ ID-σ_1(Bk)$\gg$ ID(Bk)) and the positional restrictions (recall (10): ID-σ_1(Bk) $\gg$ $^*\mathrm{F}_3$, $^*\mathrm{B}_2 \gg$ ID(Bk)) are successfully reproduced by this grammar. Note that even though the coincidental vowel harmony constraints are still highly ranked, they are not problematic because both $^*\mathrm{B}_2$ and $^*\mathrm{F}_3$ are situated correctly to catch any restricted non-initial vowels even in harmonic contexts; see Tableau (17).

(17)

/y..ø/	$^*\underline{\mathrm{B}}_5\mathrm{F}_3$	$^*\underline{\mathrm{B}}_5\ldots\mathrm{F}_3$	ID-σ_1(Bk)	$^*\mathrm{F}_3$	ID(Bk)
☞ a. y..ø				*	
b. y..o				*	*!
c. u..ø	*!	*	*	*	*
d. u..o			*!	**	

The only remaining issue here is that the difference in ranking value between $^*\mathrm{F}_3$ and ID(Bk) is relatively small, such that noise during evaluation results in the odd swapped ranking.[5] All things considered, however, Learner B demonstrates great success at learning the target North Estonian grammar.

5 Conclusion

Applying a $\mathrm{M}_{\mathrm{gen}} \gg \mathrm{M}_{\mathrm{spec}}$ bias to a learner's initial ranking values gives more general markedness constraints the opportunity to remain active in a grammar being acquired from positive learning inputs. Without such a bias - if all markedness constraints start with the same value - it is possible that less-often violated, more specific markedness constraints get credit for generating the attested patterns when more-often violated, more general constraints should in fact be the active ones. North Estonian Learners A vs B provide a clear example of this pair of differing outcomes.

Of course, there are some situations in which the general constraints are truly too general to capture the patterns of a language. In these cases the errors caused by satisfying those general

[5] Though there is not room for it here, interested readers will find analysis of and a proposed solution to this challenge in Vesik (in prep).

constraints trigger updates that demote them, allowing the more specific options to take precedence. The fact that a general-over-specific relationship is reversible given particular learning data is key to the success of this bias.

The combination of all of the biases introduced in this paper - the low-faithfulness bias, *a priori* $F_{spec} \gg F_{gen}$ bias, tempered promotion rate, and $M_{gen} \gg M_{spec}$ bias - produces a learner that is able to acquire an exceptionally successful grammar for North Estonian.

Considering the task of gradual, error-driven learning more broadly, the application of an $M_{gen} \gg M_{spec}$ bias is the most novel of the parameters discussed herein. In this particular typology, it appears to be the most useful for learning the positional restrictions of North Estonian, whereas other language patterns benefit from other combinations of biases (see Vesik (in prep) for treatment of other languages in the typology). Although this is itself a very specific learning scenario, I propose that it would be useful to include general markedness as a counterpart to specific faithfulness in a toolbox of biases for restrictive learning in general. The overlapping violation profiles encountered here are not specific to this learning environment. In this case they are explicitly built in to the constraint set, but they can also occur by accident, which is more likely to happen as constraint sets grow. This is something to consider in relation to building analyses and simulating acquisition: it is often the case that one phenomenon is addressed at a time, in one language at a time. This makes for small, neat, and tidy constraint sets. But if we broaden our view to consider multiple phenomena, or several languages in a typology, then it is very likely that constraint sets will need to grow as well, which increases the likelihood of credit problems. This bias is a tool very well-suited to mitigating such problems.

References

Albright, Adam & Bruce Hayes (2006). Modelling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations. Fanselow, Gisbert, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in Grammar*, Oxford University Press.

Angluin, Dana (1980). Inductive inference of formal languages from positive data. *Information and Control* 45, 117–135.

Ariste, Paul (1968). *A grammar of the Votic language*. Indiana University.

Baker, C.L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry* 10, 533–581.

Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.) (2022). *The Oxford Guide to the Uralic Languages*. Oxford University Press.

Boersma, Paul (1997). How we learn variation, optionality and probability. van Son, Rob (ed.), *Proceedings of the Institute of Phonetic Sciences (IFA) 21*, University of Amsterdam, Institute of Phonetic Sciences, 43–58.

Boersma, Paul (1998). *Functional phonology*. Ph.D. thesis, University of Amsterdam.

Boersma, Paul & Bruce Hayes (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:1, 45–86.

Campbell, George L., Gareth King & ProQuest (Firm) (2013). *Compendium of the world's languages*. Routledge, 3 edn.

Clements, George N. (1976). The autosegmental treatment of vowel harmony. Dressler, Wolfgang U. & Oskar E. Pfeiffer (eds.), *Phonologica 1976*, Institut für Sprachwissenschaft, Innsbruck, p. 111–119.

Dresher, Bezalel E. (1999). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30:1, 27–67.

Fejes, László, Péter Siptár & Robert M. Vago (2024). Vowel harmony in Uralic languages. van der Hulst, Harry & Nancy A. Ritter (eds.), *The Oxford Handbook of Vowel Harmony*, Oxford University Press, 840–863, 1 edn.

Gnanadesikan, Amalia E. (1995). Markedness and faithfulness constraints in child phonology. URL `http://roa.rutgers.edu`. ROA-67.

Goldsmith, John A. (1976). *Autosegmental phonology*. Ph.D. thesis, Massachusetts Institute of Technology.

Grünthal, Riho (2015). *Vepsän kielioppi [Veps Grammar]*. Suomalais-Ugrilainen Seura, Helsinki.

Grünthal, Riho (2022). Veps. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 291–307.

Hayes, Bruce (2004). *Phonological acquisition in Optimality Theory: the early stages*, Constraints in Phonological Acquisition, Cambridge University Press, 158–203.

Hayes, Bruce (2013). *OTSoft: Constraint Ranking Software*. URL `https://brucehayes.org/otsoft/OTSoftManual_2.5_April_2021.pdf`.

Hayes, Bruce, Bruce Tesar & Kie Zuraw (2013). OTSoft [software package]. URL `http://www.linguistics.ucla.edu/people/hayes/otsoft/`. Version 2.6 beta.

Jesney, Karen & Anne-Michelle Tessier (2011). Biases in harmonic grammar: the road to restrictive learning. *Natural Language & Linguistic Theory* 29, 251–290.

Karlsson, Fred (2018). *Finnish: A comprehensive grammar*. Taylor and Francis.

Kiparsky, Paul & Karl Pajusalu (2003). Towards a typology of disharmony. *Linguistic Review* 20, 217–241.

Laakso, Johanna (2022). Livonian. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 380–391.

de Lacy, Paul (2002). *The formal expression of markedness*. Ph.D. thesis, University of Massachusetts, Amherst, URL `http://roa.rutgers.edu`. ROA-542.

Léonard, Jean Léo (1993). Analysis of variation within a typological parameter: Kalevi Wiik's approach to vowel harmony in southern and eastern Balto-Finnic. *Proceedings of the VIIIth International Conference on Dialectology*.

Magri, Giorgio (2012). Convergence of error-driven ranking algorithms. *Phonology* 29:2, 213–269.

Magri, Giorgio & René Kager (2015). How to choose successful losers in error-driven phonotactic learning. Kuhlmann, Marco, Makoto Kanazawa & Gregory M. Kobele (eds.), *Proceedings of the 14th Meeting on the Mathematics of Language (MoL 2015)*, Association for Computational Linguistics, Chicago, USA, 126–138, URL `https://aclanthology.org/W15-2311/`.

Markus, Elena & Fedor Rozhanskiy (2022a). Ingrian. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 308–329.

Markus, Elena & Fedor Rozhanskiy (2022b). Votic. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 330–349.

Metslang, Helle (2022). North and standard Estonian. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 350–366.

Nikolaev, Dmitry (2018). The database of Eurasian phonological inventories: a research tool for distributional phonological typology. *Linguistics Vanguard* 4:1.

Nikolaev, Dmitry (2019). Liv sound inventory. Moran, Steven & Daniel McCloy (eds.), *PHOIBLE 2.0*, Max Planck Institute for the Science of Human History, Jena, URL `http://phoible.org/inventories/view/2448`.

Pajusalu, Karl (2022). Seto South Estonian. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 367–379.

Prince, Alan (1997a). Paninian relations. Colloquium Talk, University of Massachusetts, Amherst.

Prince, Alan (1997b). Stringency and anti-paninian hierarchies. Handout from LSA Institute.

Prince, Alan (1997c). Topics in OT: Class 3: Harmonic completeness, AP order; chain shifts. Handout from LSA Institute.

Prince, Alan (1998). Two lectures on optimality theory. Phonology Forum 1998, Kobe University.

Prince, Alan (1999). Paninian relations. Handout, University of Marburg.

Prince, Alan (2002). Entailed ranking arguments. URL `http://roa.rutgers.edu`. ROA-500.

Prince, Alan & Bruce Tesar (2004). Learning phonotactic distributions. Kager, René, Joe Pater & Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, Cambridge University Press, p. 245–291.

Pulleyblank, Douglas (2002). Harmony drivers: No disagreement allowed. Chew, Patrick (ed.), *Proceedings of the Twenty-eighth Annual Meeting of the Berkeley Linguistics Society*, BLS, 249–267.

Sang, Joel (2009). Ühest fonotaktilisest kollisioonist (Kihnu näitel) [On a phonotactic collision (on the example of Kihnu subdialect)]. *Keel Ja Kirjandus* LII:11, 809–817.

Sarhimaa, Anneli (2022). Karelian. Bakró-Nagy, Marianne, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*, Oxford University Press, 269–290.

Smolensky, Paul (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* 21, p. 720–731.

Suomi, Kari, Juhani Toivanen & Riikka Ylitalo (2008). *Finnish Sound Structure: Phonetics, phonology, phonotactics and prosody*. Studia Humaniora Ouluensia 9, University of Oulu, Oulu.

Tesar, Bruce & Paul Smolensky (1998). Learnability in optimality theory. *Linguistic Inquiry* 29:2, 229–268.

Tessier, Anne-Michelle (2007). *Biases and Stages in Phonological Acquisition*. Ph.D. thesis, University of Massachusetts, Amherst.

Van Rossum, Guido & Fred L. Drake (2009). Python [programming language]. URL `http://www.python.org/`. Version 3.8.

Vesik, Kaili (in prep). *Gradual Error-Driven Learning and Typology of Finnic Vowel Interactions*. Ph.D. thesis, University of British Columbia.