This is Chapter 4/5 only. Refer to 20240818 dissertation outline and notes for complete outline.

Kaili's very official dissertation draft

For purposes of graduating in the near-ish future

4/5. Learning biases and optimal learning conditions

As per usual, my initial work will be done in Google Docs while the final draft(s) will be in LaTeX. So any formatting issues or nuances will be sorted out later.

Yellow highlighting indicates updates/follow-up needed.

Notes and/or specific requests for feedback highlighted in blue throughout.

Prior to this chapter, I will have introduced:

- The balto-finnic languages / vowel pattern typology
- Representative sample languages (patterns) for the above
- My constraint set & OT analysis of the sample languages
- Learning algorithms, in particular GLA-type learners
    - Definition of idempotence
    - Definition of θ-notation for ranking values
    - Explain my method of calculating "average frequency of correct results"
- My methods for simulations - python script etc
- Simulated learning data for the sample languages

4.2. F_{spec} >> F_{gen}

The constraint set that I use for this project includes only two faithfulness constraints, **ID(Bk)** and **ID(Bk)-σ_1**, 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, the ranking **ID(Bk)-σ_1** >> ***F_3** >> **ID(Bk)** bans vowels in set **F_3** in general, but permits them in initial syllables.

A specific-over-general faithfulness bias (F_{spec} >> 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). I take two slightly different approaches to this idea, detailed in Sections 4.2.1 and 4.2.2.

4.2.1. *A priori* bias

One approach to the specific-over-general faithfulness bias is to ensure that the ranking value for the specific version of the constraint is a minimum specified distance higher than that of the general version [TODO citation?]. The satisfaction of this bias is checked persistently through the learning simulation, both in the initial state and after each individual learning update.

4.2.1.1. Rationale

Maintaining a minimum difference between the ranking values of a specific-general pair of faithfulness constraints 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.

4.2.1.2. Implementation

The F_{spec} >> F_{gen} bias between any specific-general pair of faithfulness constraints can be implemented by means of an *a priori* bias that ensures $\theta_{Fspec} - \theta_{Fgen} \geq d$, 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. OTSoft (Hayes et al, 2013a) sets the default value of this difference to be *d = 20*, stating that it is “very close probabilistically to being an obligatory ranking” (Hayes et al, 2013b: 24).

In my learning simulations, I test the omission of this bias as well as a range of different *d* values: 0 (θ_{Fspec} must be no less than θ_{Fgen}), 10, 20, 30, and 40.

4.2.1.3. Simulation results - *a priori*

To demonstrate the effect of the *a priori* bias, I simulate acquisition of the three sample languages using Learner B, defined with the settings in Table 8. The selection of *d* = 20 for illustrative purposes is drawn from the OTSoft default as mentioned above. Results using learners with other values of *d* are summarized in Section 4.5.

Parameter	Setting
All basic parameters	Default
A priori bias	*d* = 20

Table 8. Parameter settings for Learner B.

With the *a priori* bias set to *d* = 20, learning simulations for all three sample languages fail to acquire the target grammars. The learner trained on North Estonian data produces a grammar that, while not correct, does have some promising characteristics. On the other hand, the learners trained on Finnish and North Seto data once again produce fully-faithful grammars. Test results are summarized in Table 9.

Language	Average frequency of correct outputs
Finnish	0.2895
North Estonian	0.7796
North Seto	0.3133
Overall	0.4608

Table 9. Summary of results from simulations with Learner B.

Finnish: Table 10 shows the final ranking values for a selection of crucial constraints, after learning from simulated Finnish data. Both faithfulness constraints have risen to the top. Id(Bk)'s distance above $*B_2$ and the relevant no-disagreement constraints is small enough that evaluation noise might cause it to swap rankings with one of its neighbours. However, in order to meet the crucial rankings

$*B_2$, $*F_3 \ldots \underline{B}_5$, $*F_3\underline{B}_5$, $*\underline{B}_5 \ldots F_3$, $*\underline{B}_5 F_3$ >> Id(Bk)σ_1 >> Id(Bk)

proposed in Section 3.X, such swaps would have to be guaranteed to occur at every evaluation, which is extremely unlikely given the final ranking values. Hence the final grammar produced by Learner B on Finnish inputs is more or less fully faithful, with a representative evaluation shown in Tableau (1).

Constraint	Final ranking value
Id(Bk)σ_1 Id(Bk) $*\underline{B}_5 \ldots F_3$ $*B_2$ $*F_3\underline{B}_5$ $*F_3 \ldots \underline{B}_5$ $*\underline{B}_5 F_3$	136.000 116.000 112.000 110.000 110.000 110.000 110.000

…	…

Table 10. Excerpt of final ranking values for Finnish after simulation with Learner B.

(1) Sample evaluation of input /o..æ/ in the Finnish grammar acquired by Learner B. The grammar selects the faithful candidate [o..æ] as optimal even though it is not harmonic.

	/o..æ/	Id(Bk)σ_1	Id(Bk)	$*\underline{B}_5\ldots F_3$	… (constraints such as $*B_2$, $*F_3\underline{B}_5$, $*F_3\ldots\underline{B}_5$, $*\underline{B}_5F_3$, and others)
☞	o..æ			*	
	o..ɑ		*!		
	ø..æ	*!	*		
	ø..ɑ	*!	**		

[TODO possibly adjust the following paragraph / table after 20241128 discussion with AMT]

In theory it should have been reasonable for Id(Bk)σ_1 to end up with a final ranking value greater than or equal to the top-ranked markedness constraints with Id(Bk) lower down. However, at the time that Id(Bk)σ_1 approaches the highest-ranked markedness constraints (including $*B_2$ with θ=110), the other context-free markedness constraints all have values in [100, 106] and are therefore within a small enough window for evaluation noise to make (e.g.) $*B_3$ or $*F_3$ active in selecting the optimal candidate (see Table 10.1). This results in errors and therefore more updates which push the faithfulness constraints ever higher. It is only once Id(Bk) surpasses this clump of markedness constraints that errors taper off and the learner converges.

Trial number 69 i..e..ɑ ~ ɯ..ɤ..ɑ		Trial number 73 u..u..ɑ ~ u..y..æ		Trial number 90 u..y..ø ~ i..u..o	
$*B_2$	110	Id(Bk)σ_1	112	Id(Bk)σ_1	120
Id(Bk)σ_1	108	$*B_2$	110	$*B_2$	110
$*B_3$	106	$*B_3$	108	$*\underline{B}_5\ldots F_3$	108
$*F_3$	104	$*\underline{B}_5\ldots F_3$	106	$*B_3$	106
$*B_1$	104	$*F_3$	104	$*F_3\underline{B}_5$	106
$*\underline{B}_5\ldots F_3$	104	$*B_1$	104	$*F_3\ldots\underline{B}_5$	106
$*F_1$	102	$*F_3\ldots\underline{B}_3$	104	$*\underline{B}_5F_3$	106
$*F_3\underline{B}_3$	102	$*F_3\underline{B}_5$	104	$*F_3$	104
$*F_3\ldots\underline{B}_3$	102	$*F_3\ldots\underline{B}_5$	104	$*B_1$	104
$*F_3\underline{B}_5$	102	$*\underline{B}_5F_3$	104	$*F_3\ldots\underline{B}_3$	104
$*F_3\ldots\underline{B}_5$	102	$*F_3\underline{B}_3$	102	$*\underline{B}_5F_1$	104
$*\underline{B}_3\ldots F_3$	102	$*F_4\ldots\underline{B}_3$	102	$*\underline{B}_5\ldots F_1$	104
$*\underline{B}_5F_1$	102	$*\underline{B}_3\ldots F_3$	102	$*F_1$	102

$*\underline{B}_5 \ldots F_1$	102	$*\underline{B}_5 F_1$	102	$*F_3\underline{B}_3$	102
$*\underline{B}_5 F_3$	102	$*\underline{B}_5 \ldots F_1$	102	$*\underline{B}_3 \ldots F_3$	102
$*F_4$	100	$*F_1$	100	Id(Bk)	100
$*F_5$	100	$*F_4$	100	$*F_4$	100
$*B_5$	100	$*F_5$	100	$*F_5$	100
…		$*B_5$	100	$*B_5$	100
Id(Bk)	88	…		…	
		Id(Bk)	92		

Table 10.1. The highest of Finnish Learner B's constraint ranking values after three different learning updates. Although the crucial constraints (ideally $*B_2$ >> Id(Bk)σ_1 >> $*\underline{B}_5 \ldots F_3$, $*F_3\underline{B}_5$, $*F_3 \ldots \underline{B}_5$, $*\underline{B}_5 F_3$) are in reasonably good positions, constraints such as $*B_3$, $*B_5$, $*F_1$, $*F_3$, $*F_4$, $*F_5$ are near enough to be potentially disruptive.

North Estonian: The final ranking values for a selection of crucial constraints, after learning from simulated North Estonian inputs, are shown in Table 11. Several of the crucial relative rankings $*B_1$ >> Id(Bk)σ_1 >> $*F_3$, $*B_2$ >> Id(Bk), proposed in Section 3.X, are met by this grammar. However, one of the key elements – the full ban on vowels from set B_1 – is missing, by virtue of that fact that $*B_1$'s final value is not only not at the top, but below even Id(Bk). Thus the acquired grammar will incorrectly permit B_1 vowels in initial syllables.

Constraint	Final ranking value
Id(Bk)σ_1	129.220
$*B_2$	115.000
…	…
$*F_3$	110.220
Id(Bk)	109.220
$*B_1$	108.000
$*\underline{B}_5 \ldots F_3$	108.000
…	…
$*F_5 \ldots \underline{B}_2$	106.000
…	…
$*F_5\underline{B}_2$	104.600
…	…
$*\underline{B}_5 F_3$	104.000
…	…

Table 11. Excerpt of final ranking values for North Estonian after simulation with Learner B. A selection of no-disagreement constraints are included here for the purpose of comparison with results of simulations discussed in subsequent sections.

The ranking acquired by this learner does generally follow the required positional restrictions by ranking Id(Bk)σ_1 >> $*F_3$, $*B_2$ >> Id(Bk); however, the ranking values are close enough together that the stochastic nature of evaluation results in somewhat variable adherence to these positional restrictions. For example, ungrammatical test input /y..æ/ would be expected to surface as [y..ɑ], neutralizing the restricted vowel in the second syllable. However, during testing, this grammar selects output candidates with the frequencies shown in Table 12.

/y..æ/	Output frequency
y..æ	0.35
y..ɑ	0.65
u..æ	0.00
u..ɑ	0.00

Table 12. Frequency of candidate selection for input /y..æ/ with North Estonian grammar acquired by Learner B. Number of sample evaluations = 100.

Although there is some crowding, at least one success the North Estonian learner achieves that the Finnish one does not is that the learner converges with several markedness constraints between the two faithfulness constraints. [TODO add some explanation here after 20241128 discussion with AMT]

North Seto: Table 13 shows the final ranking values for a selection of crucial constraints, after learning from simulated North Seto data. Similar to the Finnish results, both faithfulness constraints have risen to the top. Though the markedness constraints are correctly ordered relative to each other, the relative positions of the faithfulness vs the markedness constraints are not correct with respect to the crucial rankings proposed in Section 3.X:

$*F_4\ldots\underline{B}_5$, $*F_4\underline{B}_5$, $*\underline{B}_5\ldots F_4$, $*\underline{B}_5F_4$, Id(Bk)syl1 >> *B1 >> Id(Bk)

Again, the final grammar produced by Learner B on North Seto inputs is essentially fully faithful, with similar learning challenges as described for Finnish.

Constraint	Final ranking value
Id(Bk)σ_1	136.000
Id(Bk)	116.000
$*F_4\ldots\underline{B}_5$	110.000
$*\underline{B}_5\ldots F_4$	110.000
$*F_4\underline{B}_5$	108.000
$*\underline{B}_5F_4$	108.000
…	…
$*B_1$	104.000
…	…

Table 13. Excerpt of final ranking values for North Seto after simulation with Learner B.

Results from both Section 4.2.1 and Section 4.2.2 are discussed in Section 4.2.4.