Nat Lang Linguist Theory (2011) 29: 251–290
DOI 10.1007/s11049-010-9104-2

# Biases in Harmonic Grammar: the road to restrictive learning

**Karen Jesney · Anne-Michelle Tessier**

Received: 25 September 2009 / Accepted: 2 November 2009 / Published online: 27 July 2010

**Abstract** In the Optimality-Theoretic learnability and acquisition literature it has been proposed that certain classes of constraints must be biased toward particular rankings (e.g., Markedness $\gg$ IO-Faithfulness; Specific IO-Faithfulness $\gg$ General IO-Faithfulness). While sometimes difficult to implement efficiently or comprehensively, these biases are necessary to explain how learners acquire the most restrictive grammar consistent with positive evidence from the target language, and how innovative patterns emerge during the course of child phonological development. This paper demonstrates that altering the mode of constraint interaction from strict ranking as in Optimality Theory to additive weighting as in Harmonic Grammar (HG) reduces the number of classes of constraints that must be distinguished by such biases. Using weighted constraints and a version of the Gradual Learning Algorithm (GLA), the only distinction needed is between Output-based constraints, which must be biased toward high weights, and Input-Output-based constraints, which must be biased toward the lowest weights possible. We implement this distinction within the HG-GLA model by assigning different initial weights and plasticity values to the two classes of constraints. This implementation suffices to ensure that restrictive grammars are

---

We gratefully acknowledge the extensive help of Joe Pater on every stage of this work, as well as the contributions of Michael Becker, Paul Boersma, Gaja Jarosz, John McCarthy, Pat Pratt, Bruce Tesar, Matt Wolf, and the audiences at NELS38 and NecPhon1. We also wish to thank the three anonymous *NLLT* reviewers and the associate editor for their very thorough, challenging, and insightful comments. This work has been supported in part by NSF grant BCS-0813829 to the University of Massachusetts Amherst, and by SSHRC Doctoral Fellowship #752-2005-1708 to K. Jesney.

K. Jesney
226 South College, Department of Linguistics, University of Massachusetts Amherst, Amherst, MA 01003, USA
e-mail: kjesney@linguist.umass.edu

A.-M. Tessier (✉)
University of Alberta, 4-22 Assiniboia Hall, Edmonton, Alberta, T6G 2E7, Canada
e-mail: amtessier@ualberta.ca

learned, and also predicts the emergence of a variety of attested intermediate stages during the course of acquisition.

**Keywords** Phonological acquisition · Learnability theory · Biases · Harmonic Grammar · Optimality Theory · Gradual Learning Algorithm

## 1 Introduction

In the Optimality-Theoretic learnability and acquisition literature it has been frequently argued that certain classes of constraints must be biased toward particular rankings. These biases have two primary motivations: ensuring that restrictive final grammars can be successfully reached and allowing the intermediate stages attested in acquisition to emerge. Biases distinguishing between at least four classes of constraints have been proposed; these are given in (1).

(1) *Four classes of constraints must be distinguished for restrictiveness in OT learning:*
   a. **Markedness** constraints should dominate **IO-Faithfulness** constraints. (Demuth 1995; Gnanadesikan 2004; Pater 1997; Smolensky 1996)
   b. **Specific IO-Faithfulness** constraints should dominate **General IO-Faithfulness** constraints. (Hayes 2004; Prince and Tesar 2004; Smith 2000; Tessier 2007)
   c. **OO-Faithfulness** constraints should dominate **Markedness** constraints. (Hayes 2004; Tessier 2006, 2007; see also McCarthy 1998 for arguments that OO-Faithfulness should dominate IO-Faithfulness)

While intuitively simple, biases of this sort have proven challenging to implement. In this paper we show that altering the mode of constraint interaction from strict ranking as in Optimality Theory (Prince and Smolensky 1993; McCarthy and Prince 1995) to additive weighting as in Harmonic Grammar (Legendre et al. 1990; Smolensky and Legendre 2006; Pater 2009; Potts et al. 2010) substantially reduces the number of constraint classes that must be distinguished in order to ensure restrictive learning. Using weighted constraints and a simple GLA-style learning algorithm (Boersma 1998; Boersma and Hayes 2001; Boersma and Levelt 2003; Boersma and Pater 2007, 2008) the four constraint classes in (1) can be reduced to the two constraint classes in (2).[1]

(2) *Only two classes of constraint must be distinguished for restrictiveness in HG learning:*
   a. **Output-based** constraints should be weighted high—high enough to remain inviolable if necessary.
   b. **Input-Output-based** constraints should be weighted as low as possible.

---

[1] The implementation of high and low weightings for these two constraint classes, and the predictions that result, will be a central focus of the rest of this paper.

The class of OUTPUT-BASED CONSTRAINTS includes Markedness and OO-Faithfulness. Constraints in this class are unified in two ways: they are able to force deviations from the input form, and they can be assessed solely with reference to output forms generated by the grammar. Restrictiveness in HG requires that constraints in this class be biased toward high weights. INPUT-OUTPUT-BASED CONSTRAINTS include both General and Specific IO-Faithfulness constraints; restrictiveness requires that these constraints be biased toward weights that are as low as possible. We implement these biases in the HG-GLA model developed here by assigning an initial weight of zero—the lowest possible weight in our model—to Input-Output-based constraints and a high positive initial weight to Output-based constraints. Additionally, we require that the weights of Input-Output-based constraints be adjusted more slowly than those of Output-based constraints, thereby encouraging persistent low weights for Input-Output-based constraints. We demonstrate for a variety of cases that this is sufficient to ensure that the final grammar acquired by the gradual learner is restrictive, generating only forms consistent with the target language. Furthermore, this approach predicts that a range of attested intermediate stages will naturally emerge during the course of acquisition, based solely on positive evidence from the target language.

Section 2 presents the model of weighted constraint interaction employed in this paper and introduces requirements for restrictiveness in Harmonic Grammar, focusing on the need for distinct treatments of Output-based constraints and Input-Output-based constraints in this framework. Section 3 describes the HG-GLA learner we adopt and discusses the means by which we implement the distinctions between the two classes of constraints, illustrating with examples that use Markedness and IO-Faithfulness. Section 4 demonstrates how this HG-GLA model predicts the emergence of attested intermediate stages in L1 acquisition, while Sect. 5 compares the HG-GLA model to previous proposals in the Optimality-Theoretic learnability and acquisition literature. Section 6 concludes that the HG-GLA is a viable model, offering certain benefits relative to ranked-constraint approaches.

## 2 Harmonic Grammar

### 2.1 The grammar model

Like Optimality-Theoretic grammars, Harmonic Grammars have three components: GEN, a function that takes an input form and returns a series of output candidates; CON, the set of constraints (often taken to be universal); and EVAL, the means of constraint interaction that serves to select an optimal output candidate from among the set provided by GEN. The first two components—GEN and CON—are in principle the same in OT and HG; it is only the mode of constraint interaction—EVAL—that differs between the two theories.[2]

---

[2] Adopting a model of weighted constraint interaction may well lead to new hypotheses about the nature of GEN and the contents of CON. A number of recent papers, for example, have explored the possibility of weighted constraint grammars using positively-formulated constraints and/or language-specific constraints induced from the observed surface patterns of the target language (e.g., Boersma and Pater 2007; Hayes

In OT, constraints are strictly ranked. Candidates are assessed based on this hierarchy in a manner such that violations of lower-ranking constraints are relevant only when higher-ranking constraints alone cannot select a single optimum. The OT tableau in (3) illustrates:

(3) *OT strict domination*

| | | | C1 | C2 | C3 |
|---|---|---|---|---|---|
| i. | | [A] | *! | | |
| ii. | ☞ | [B] | | * | * |
| iii. | | [C] | | **! | |

Here, the highest-ranked constraint—C1—rules out candidate (3.i), which performs worse on this constraint than either of the other candidates. The decision between candidates (3.ii) and (3.iii) is then passed down to the next constraint—C2. These two candidates both violate C2, but (3.ii) does so to a lesser extent than (3.iii), incurring one violation rather than two. Candidate (3.ii) is thus deemed optimal. According to the strict-domination logic of OT, the fact that (3.i) does not violate C2 or C3 at all is irrelevant; (3.i) is ruled out by C1, the highest-ranking constraint, and no virtues it may display with respect to lower-ranking constraints can redeem it. The gray-shading of cells reflects this logic. In tableau (3) the cells under C3 are all grayed out, for example, reflecting the fact that C1 and C2 have already selected an optima from among the three candidates. (See McCarthy 2002 for thorough discussion of OT's strict-domination approach to candidate evaluation.)

In Harmonic Grammar, constraints are weighted rather than ranked and EVAL operates by selecting the candidate with the highest numeric Harmony value. Each constraint—C1, C2, C3, …, Cn assigns to each candidate a violation score—$v_1, v_2, v_3, \ldots, v_n$; this violation score is a negative number corresponding to the number of violations of that constraint. A candidate $R$'s Harmony ($H$) is determined by multiplying each violation score by the corresponding constraint's weight—$w_1, w_2, w_3, \ldots, w_n$, and then summing as in (4).

(4) $$H(R) = v_1(R)w_1 + v_2(R)w_2 + v_3(R)w_3 + \cdots + v_n(R)w_n$$

Following Keller (2000, 2006), we diverge from the original HG proposal by limiting constraint weights to non-negative real numbers (see also Pater et al. 2007; Prince 2003; and Potts et al. 2010). Since constraint violations are treated as penalties (negative numbers), the highest possible Harmony value for any candidate is zero, making the optimum the candidate whose Harmony value is closest to zero.

The HG tableau in (5) provides an illustration. The weight of each constraint is shown immediately below that constraint's name; in (5), C1 has a weight of 5, C2 has a weight of 3, and C3 has a weight of 1. The rightmost column calculates the Harmony

---

and Wilson 2008; Wilson 2006). Our discussion focuses on broad classes of constraints (Output-based and Input-Output-based) that will likely prove necessary in any constraint-based model of phonology. We thus take the findings in this paper to be a general argument for a weighted constraints approach, independent of the specific constraints employed, or the means by which the constraints become part of the learner's system.

value of each candidate. The winning candidate—i.e., the one that has the highest Harmony—is indicated with the symbol ☞. In (5) the winning candidate is (5.ii), which wins with a Harmony value of −4, violating C2 and C3 each once; the losing candidates (5.i) and (5.iii) have lower Harmony values, −5 and −6, respectively.

(5) *Basic HG additive interaction*

| | | C1<br>$w = 5$ | C2<br>$w = 3$ | C3<br>$w = 1$ | *H* |
|---|---|---|---|---|---|
| i. | [A] | −1 | | | $-1 \times 5 = \mathbf{-5}$ |
| ii. | ☞ [B] | | −1 | −1 | $-1 \times 3 + -1 \times 1 = \mathbf{-4}$ |
| iii. | [C] | | −2 | | $-2 \times 3 = \mathbf{-6}$ |

While the result in (5) is very similar to what would occur within a ranked constraint system where C1 ≫ C2 ≫ C3 as in (3) above, this need not be the case. The additive property of HG allows multiple lower-weighted constraint violations to *gang up* to overcome violations of a higher-weighted constraint. Indeed, most of the results discussed in this paper rely upon this type of cumulative constraint interaction. The basic type of situation of interest here is schematized in (6), where all that has been altered from (5) are the weights of C1 (now 4 rather than 5) and C3 (now 2 rather than 1).

(6) *HG additive interaction with gang effect*

| | | C1<br>$w = 4$ | C2<br>$w = 3$ | C3<br>$w = 2$ | *H* |
|---|---|---|---|---|---|
| i. | ☞ [A] | −1 | | | $-1 \times 4 = \mathbf{-4}$ |
| ii. | [B] | | −1 | −1 | $-1 \times 3 + -1 \times 2 = \mathbf{-5}$ |
| iii. | [C] | | −2 | | $-2 \times 3 = \mathbf{-6}$ |

Candidate (6.i) violates only C1, and so has a Harmony value of −4. Candidate (6.ii) violates C2 and C3, giving it a Harmony value of −5, while candidate (6.iii) violates C2 twice, giving it a Harmony value of −6. The Harmony value of (6.i) is thus greater than the Harmony value of either (6.ii) or (6.iii), and so (6.i) is selected as optimal. It is the *weighting conditions*—i.e., the relationships between the constraint weights—that are key to this result, not the precise numeric constraint weights *per se*. The pattern in (6) emerges because the weight of C1 is less than the summed weight of C2 and C3, making (6.i) more harmonic than (6.ii), and because the weight of C1 is less than two times the weight of C2, making (6.i) more harmonic than (6.iii).

Crucially, the additive mode of evaluation in HG allows *all* constraint violations, including those of low-weighted constraints, to affect the selection of an optimal output form. Cells in HG tableaux can never be shaded, unlike in the OT tableau seen earlier, because *all* violation scores are relevant to the determination of an HG optimum (compare (3) to (6)). This is the primary difference between HG and OT modes of EVAL, and plays a crucial role in defining the conditions on restrictiveness in Harmonic Grammar.

## 2.2 Restrictiveness in Harmonic Grammar

More than one set of weighting conditions is often compatible with the observed output forms of a language. Typically, however, some of these sets of weighting conditions will also generate a range of structures that are *not* observed in the target language. The challenge is to find the set of weighting conditions that generates the types of structures observed in the target *and as few unobserved ones as possible*—this is the most ‘restrictive’ grammar possible, given the target language data. In this section we consider three patterns that provide insight into the conditions associated with restrictiveness in HG, and clarify the requirements for ensuring that the grammar does not overgenerate.

The first pattern, and the simplest illustration of the restrictiveness problem, is found in the interaction of conflicting Output-based Markedness constraints and Input-Output-based IO-Faithfulness constraints. In the most basic case, where violations of Markedness and IO-Faithfulness trade against each other in a one-to-one fashion, the complete illegality of a marked structure in the target language is captured in HG if the weight of the Markedness constraint is greater than the weight of the conflicting IO-Faithfulness constraint. This is essentially the same as the M $\gg$ F condition that ensures restrictiveness in OT (e.g., Smolensky 1996).

To illustrate, we consider the English ban on onsets composed of a coronal stop followed by a lateral approximant (for discussion see, e.g., Moreton 2000, 2002). This prohibition means that input /tl/ consistently maps unfaithfully. In a simplified system with only two possible forms—/kl/ and /tl/—both of these will map to [kl]; output [tl] should never prove optimal. The conflicting constraints in this case are the Markedness constraint *TL, which assigns violation marks to onset [tl] sequences, and the IO-Faithfulness constraint IdentPlace, which assigns violation marks to segments in the output that differ in specification for Place relative to their input correspondents. All weightings of the two constraints will generate the attested output [kl] because the mapping /kl/ → [kl] violates neither of the constraints. The weighting condition in (7) is necessary, however, if we are also to ensure that input /tl/ sequences never surface faithfully—i.e., that surface [tl] is banned.

(7) *Banning* [tl] *requires that the weight of Markedness be greater than the weight of IO-Ffaithfulness*
$w$(*TL) > $w$(IdentPlace)

When this condition is met, as in (8), the language is one that, like English, lacks coronal stop + lateral onset sequences in its output inventory.

(8) /tl/ *sequences are banned in the output when* $w(M) > w(F)$

| | | /tl/ | *TL<br>$w = 3$ | IdentPlace<br>$w = 1$ | $H$ |
|---|---|---|---|---|---|
| i. | | [tl] | $-1$ | | $-1 \times 3 = -\mathbf{3}$ |
| ii. | ☞ | [kl] | | $-1$ | $-1 \times 1 = -\mathbf{1}$ |

While the HG weighting conditions and the OT ranking conditions that lead to restrictiveness in this case are essentially the same—Output-based Markedness constraints must have higher weights or rankings than Input-Output-based Faithfulness constraints—the restrictiveness requirements of the two models begin to diverge when further interacting constraints are added. The rest of this section considers the consequences of including additional Output-based constraints that are violated under pressure from either Input-Output-based constraints or other Output-based constraints.

The first case adds an Output-based constraint that is readily violated in the target language under pressure from IO-Faithfulness. We illustrate this with a slight expansion of the English scenario described above, through the addition of the Markedness constraint *K which is violated by any dorsal stop in the output. While *TL is inviolable in English, *K is regularly violated and many instances of output dorsal stops are observed. Modeling the admissibility of dorsal stops is a straightforward matter of ensuring that the weight of IDENTPLACE is greater than the weight of *K. With this in place, however, guaranteeing that coronal stops are still disallowed before laterals is no longer as simple as maintaining the $w(\text{*TL}) > w(\text{IDENTPLACE})$ weighting condition discussed above. Like IDENTPLACE, *K prefers that input /tl/ sequences not map to [kl] in the output. It is thus possible for *K and IDENTPLACE to gang up and overcome the weight of *TL, resulting in an insufficiently restrictive grammar that admits coronal stops before laterals in the output.

(9) ***Non-restrictive***: /tl/ *sequences are allowed in the output when w(IDENT) + w(*K) > w(*TL)*

| | /tl/ | *TL<br>$w=3$ | IDENT<br>$w=2$ | *K<br>$w=2$ | *H* |
|---|---|---|---|---|---|
| i. | ☞ tl | −1 | | | $-1\times 3 = \mathbf{-3}$ |
| ii. | kl | | −1 | −1 | $-1\times 2 + -1\times 2 = \mathbf{-4}$ |

For the grammar to be restrictive, the weight of the inviolable Markedness constraint *TL must be greater than the *summed* weight of IDENTPLACE and the violable Markedness constraint *K. More generally, the weight of any inviolable constraint must be greater than the summed weights of all conflicting Faithfulness and Markedness constraints in order for restrictiveness to be maintained.

(10) ***Restrictive***: /tl/ *sequences are banned in the output when w(*TL) > w(IDENT) + w(*K)*

| | /tl/ | *TL<br>$w=3$ | IDENT<br>$w=1$ | *K<br>$w=1$ | *H* |
|---|---|---|---|---|---|
| i. | tl | −1 | | | $-1\times 3 = \mathbf{-3}$ |
| ii. | ☞ kl | | −1 | −1 | $-1\times 1 + -1\times 1 = \mathbf{-2}$ |

This clearly contrasts with the OT scenario. If *TL is the highest-ranked constraint in an OT grammar, strict domination ensures that a candidate that violates *TL (like (9.i)) will never be optimal compared to one that satisfies it (like (9.ii)).

The second expanded scenario also adds an Output-based constraint that is regularly violated—in this case, however, the violation is due to pressure from another Output-based constraint. We illustrate with an example based on Modern Greek. In

Modern Greek, the marked palatal segments [ç] and [j] are in complementary distribution with the velars [x] and [ɣ]; the palatals appear just before front vowels, while the velars appear in all other contexts (Kazazis 1972). This distribution is illustrated below using verbal paradigms that show this alternation. (We return to this data in Sect. 3.1 and Sect. 4.)

(11) *Greek palatal-velar alternations in the imperfect present tense*

(Kazazis 1969:384)

| “to have”—imperfect present tense | | | | “to leave”—imperfect present tense | | | |
|---|---|---|---|---|---|---|---|
| 1sg | [éxo] | 1pl | [éxume] | 1sg | [févɣo] | 1pl | [févɣume] |
| 2sg | [éçis] | 2pl | [éçete] | 2sg | [févjis] | 2pl | [févjete] |
| 3sg | [éçi] | 3pl | [éxune] | 3sg | [févji] | 3pl | [févɣune] |

As with other patterns of complementary distribution, two Markedness constraints are necessary to capture this pattern. We use a Specific Markedness constraint militating against velars before front vowels (*XI) and a General Markedness constraint disfavoring palatal segments (*Ç). These Markedness constraints conflict with each other, and with the IO-Faithfulness constraint IdentPalatal. In the properly restrictive grammar, however, IdentPalatal should not affect whether a velar or palatal continuant appears in a given context; the distribution of the two places of articulation should be determined entirely by the Markedness constraints.

In order to capture this distribution in HG, the weight of specific *XI must be greater than the weight of general *Ç, and the weight of both Markedness constraints must be greater than the weight of IdentPalatal. This alone is not necessarily adequate, however, as (12) shows. While the OT ranking that corresponds to the weighting in (12) would achieve the correct results, cumulative constraint interaction makes the HG scenario more complex.

(12) ***Non-restrictive:*** */xi/ sequences are allowed in the output when* $w(\text{*Ç}) + w(\text{Ident}) > w(\text{*XI})$

| | /xi/ | *XI<br>$w = 4$ | *Ç<br>$w = 3$ | IdentPal<br>$w = 2$ | *H* |
|---|---|---|---|---|---|
| i. | ☞ xi | $-1$ | | | $-1 \times 4 = \mathbf{-4}$ |
| ii. | çi | | $-1$ | $-1$ | $-1 \times 3 + -1 \times 2 = \mathbf{-5}$ |

Both *Ç and IdentPalatal prefer that input /xi/ sequences map faithfully to [xi] in the output; these constraints can therefore gang up to overcome the weight of *XI, resulting in an insufficiently restrictive grammar that admits velar fricatives before front vowels. The properly-restrictive grammar requires that the weight of *XI, which is never violated in the output, be greater than the *summed* weight of *Ç and IdentPalatal (13.a). This is a further case of the general requirement seen in the previous example: for restrictiveness to be ensured, the weight of an inviolable constraint must be greater than the summed weights of the conflicting Output-based and Input-Output-based constraints.

(13) *Necessary weighting conditions for restrictiveness in Modern Greek*

a. $w(\text{*XI}) > w(\text{*Ç}) + w(\text{IdentPal})$
b. $w(\text{*Ç}) > w(\text{IdentPal})$

(14) ***Restrictive****: /*xi*/ sequences are banned in the output when w(*XI) > w(*Ç) + w(*IDENT*)*

| | /xi/ | *XI $w = 4$ | *Ç $w = 2$ | IDENTPAL $w = 1$ | $H$ |
|---|---|---|---|---|---|
| i. | xi | $-1$ | | | $-1 \times 4 = \mathbf{-4}$ |
| ii. | ☞ çi | | $-1$ | $-1$ | $-1 \times 2 + -1 \times 1 = \mathbf{-3}$ |

The difference between the Modern Greek and English scenarios lies in the relative weightings of the two lower-weighted constraints. In the Modern Greek case, the weight of *Ç is greater than that of IDENTPALATAL, and so /ç/ maps faithfully only when it appears before a front vowel—i.e., when faithful mapping allows inviolable *XI to be satisfied. The distribution of the marked segment [ç] is determined by Markedness, not by Faithfulness, and so the weight of *Ç must be greater than the weight of IDENTPALATAL. This $w$(MARKEDNESS) > $w$(IO-FAITHFULNESS) configuration results in a scenario that is ultimately more restrictive than the one observed in English. English [k], unlike Modern Greek [ç], is admitted in any environment, and so the facts of the target language demand that the weight of the IO-Faithfulness constraint be greater than the weight of Markedness in this case.

While English and Modern Greek differ in the weighting conditions affecting the violable constraints, in both cases restrictiveness requires that the weights of Output-based constraints that are *never* violated in the language—*XI and *TL—be greater than the summed weight of the conflicting constraints that are violated *at least sometimes* in the language. Put more generally, restrictiveness requires that the weights of violated constraints (whether Output-based or Input-Output-based) be low enough that, even cumulatively, they do not overcome the weights of inviolable Output-based constraints. Further restrictiveness is achieved when, among the constraints that are at least sometimes violated, the weights of Output-based constraints are greater than the weights of conflicting Input-Output-based constraints. The weights of Input-Output-based constraints should therefore exceed those of Output-based constraints only when necessary to account for the observed language data, as with the English case. Our first statement of these restrictiveness conditions was given in the introduction in (2); it is repeated below in (15):

(15) *Requirements for restrictiveness in HG*

a. Output-based constraints should be weighted high—high enough to remain inviolable if necessary.

b. Input-Output-based constraints should be weighted as low as possible.

Importantly, only two classes of constraints are distinguished here: Input-Output-based constraints and Output-based constraints. It is our conclusion that this simple two-way distinction is adequate for restrictiveness in Harmonic Grammar. As Sects. 4 and 5 discuss, additional classes of constraints such as Specific IO-Faithfulness and OO-Faithfulness do not need to be distinguished in this model. First, though, Sect. 3 describes an algorithm that aims to bring (15) home to the learner while relying only on positive evidence.

## 3 Learning a restrictive Harmonic Grammar with the GLA

The statement in (15) is something of a heuristic: it captures the biases we want our final grammar to have, not how it will come to have them. If a full range of positive and negative evidence is available, finding a restrictive set of weights is a relatively straightforward matter (see Potts et al. 2010 for discussion). However, the actual task facing the language learner is rather more complex; children only encounter forms that are part of the language they are learning (positive evidence), and they encounter these forms gradually over time. Furthermore, children have no means of knowing *a priori* whether a given constraint is violated or not in the target language.

This section spells out our proposal for a restrictive HG learner that implements the biases in (15) while gradually learning on the basis of positive target language evidence. Section 3.1 begins with the initial state: the grammatical starting point that instantiates the bias in its unadulterated form, before the influence of any language-specific evidence. Section 3.2 describes the learning procedure itself: the HG-GLA (Boersma and Pater 2007, 2008; Boersma and Weenink 2009). We then demonstrate that the bias must be imposed continually throughout learning, and implement this persistent bias through differences in constraint plasticity (how and why is the province of Sect. 3.3). Finally, Sect. 3.4 draws together the results of this section, and presents the details of the learner simulations that follow in Sect. 4.

### 3.1 The biased initial state in HG

As discussed in Sect. 2.2, restrictive constraint-based grammars require that, in the absence of target language evidence to the contrary, Markedness constraints have weights or rankings higher than those of IO-Faithfulness constraints (see Smolensky 1996; Prince and Tesar 2004). Beyond coinciding with the requirements of restrictive adult grammars, this initial configuration of the constraints accords well with a wealth of evidence from children's earliest productions. By and large, early grammars tolerate much less markedness than their adult counterparts, a fact that is implemented through MARKEDNESS $\gg$ IO-FAITHFULNESS initial states in the OT literature (see, among others, Curtin and Zuraw 2002; Demuth 1995; Gnanadesikan 2004; Hayes 2004; Kehoe and Stoel-Gammon 1997; Pater 1997; Stampe 1973).[3] In this paper, we have generalized this requirement, biasing all Output-based constraints (OO-Faithfulness as well as Markedness) to have weights greater than all IO-Faithfulness constraints in the initial state.[4]

---

[3] An initial $w$(MARKEDNESS) > $w$(IO-FAITHFULNESS) bias to model this type of child data is suggested for Maximum Entropy OT—a related weighted constraint system—by Goldwater and Johnson (2003) and Jäger (2007).

[4] Hale and Reiss (1998) argue that IO-Faithfulness constraints must outrank Markedness constraints in the initial state of OT grammar learning. Under these assumptions, errors in children's early productions are attributed to performance factors rather than to the grammar *per se*. While it is true that young infants are often able to perceive contrasts that are absent from the target language (e.g., Werker and Tees 1984), it is not the case that children's underlying representations consistently match those of adults (e.g., Macken 1980; Velleman 1988). Furthermore, it has been shown that constraint-based learners who can perceive faithfully need not have equally-faithful production grammars (Smolensky 1996; Tesar 1997; Tesar and Smolensky 1998, 2000; Pater 2004).

A simple illustration of the effect of an initial $w$(Output-based) > $w$(Input-Output-based) bias comes from a target language that includes a particularly marked segment in its inventory—for example, the Modern Greek palatal fricative [ç]. This is a relatively uncommon segment cross-linguistically (Ladefoged and Maddieson 1996), and it is highly unlikely that a child acquiring such a language would include it within his or her initial inventory, especially in contexts where it precedes a non-front vowel. According a higher initial weight to the Markedness constraint *Ç than to the IO-Faithfulness constraint IdentPalatal is a first step in capturing this dispreference.

As Sect. 2.2 demonstrated, however, simply assigning higher weights to Markedness constraints than to IO-Faithfulness constraints is inadequate to ensure that the marked segment will be excluded in HG. The possibility of cumulative constraint interaction makes it necessary to define the weighting conditions more precisely. Our first question is thus: in the initial state, what should the constraint weights be?

There are two components to this answer, both of which come from the statement of restrictiveness given in (15). To ensure that the grammar does not overgenerate, Input-Output-based constraints must be weighted as low as possible. In the current model, where constraint weights are limited to non-negative real numbers, this lowest possible weight is zero. Output-based constraints, for their part, must have high weights. Because there is no upper-limit on constraint weights, 'high' cannot be precisely defined. The specific value required in order to ensure that Output-based constraints that are never violated in the target language remain inviolable will depend on a variety of factors, including the number and nature of the constraints that are hypothesized to be part of Con. For the cases discussed in this paper, however, an initial weight of 100 for Output-based constraints is adequate.

(16) Initial state: $w$(Output-oriented) e.g., $w$(*Ç), $w$(*XI): 100
$w$(Input-Output-oriented) e.g., $w$(IdentPalatal): 0

Assuming that learners begin by attempting to faithfully reproduce observed surface forms (cf. Smith 1973; Hayes 2004; Pater 1997),[5] the Modern Greek learner will have to contend with input forms like /çéri/ 'hand' that contain palatal fricatives. As (17) shows, the initial state from (16) will result in errors that place a priority on Output-based (Markedness) pressures, as is generally seen in child language. The target output is not consistently selected by this initial-state grammar; instead the two candidates in this tableau are tied, with equal Harmony values. Faced with multiple equally harmonic outputs, the speaker randomly selects and produces one of the winners. As a result, the non-target-like velar form, i.e., /çéri/ → *[xéri] (17.ii), will be selected as optimal with .5 probability.[6]

---

[5] We abstract away, here, from various complicating details; not least of all, what learners observe of surface forms is only an acoustic signal, and not a full phonological representation (see Tesar 1997; Boersma 1998; and Tesar and Smolensky 1998, 2000 for discussion).

[6] Application of noise to the constraint weights as in stochastic OT (Boersma 1998; Boersma and Hayes 2001) is an equally effective means by which the learner might choose between outputs in (17), and may well have additional benefits for learning with more complex systems (Boersma and Pater 2008). For the relatively simple cases discussed here, the random selection approach is adequate. For learning paths similar to those in Sect. 4 simulated using the noisy HG-GLA, see Jesney and Tessier (2009).

(17) *Initial-state treatment of Greek fricatives: Markedness determines the realization of /ç/*

| | /çéri/ | *Ç $w = 100$ | *XI $w = 100$ | IdentPal $w = 0$ | $H$ |
|---|---|---|---|---|---|
| i. | ☞ çéri | $-1$ | | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ xéri | | $-1$ | $-1$ | $-1 \times 100 + -1 \times 0 = \mathbf{-100}$ |

Choosing the non-target-like output in (17) will lead us to learning in Sect. 3.2 below. This relies, however, on the initial-state grammar selecting an optimum based solely on Output-based considerations. In (17), the IO-Faithfulness constraint IdentPalatal has an initial weight of zero and therefore cannot affect the outcome in any way. The situation would be rather different if IO-Faithfulness did not begin with the lowest possible weight. If IdentPal were weighted even at 1, it would be able to gang up with *XI and cause the target palatal form to consistently surface.[7]

(18) *Consequence of IdentPal weighted at 1: no errors*

| | /çéri/ | *Ç $w = 100$ | *XI $w = 100$ | IdentPal $w = 1$ | $H$ |
|---|---|---|---|---|---|
| i. | ☞ çéri | $-1$ | | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | xéri | | $-1$ | $-1$ | $-1 \times 100 + -1 \times 1 = \mathbf{-101}$ |

Given this alternative initial state where Input-Output-based constraints do not have the lowest possible weight, the correct (but marked) output [çéri] is selected from the very beginning. This is an undesirable result for two reasons. First, the child's grammar is consistently producing output forms with the marked segment [ç], contrary to the patterns generally observed in acquisition. Second, in the absence of errors, no changes to the initial constraint weights are motivated. With the constraint weights remaining unchanged, the adult grammar is clearly unrestrictive, allowing velar fricatives to surface faithfully before front vowels, breaking the target pattern of complementary distribution.

(19) *Without learning: final grammar **unrestrictive** in the face of illicit velars*

| | /xéri/ | *Ç $w = 100$ | *XI $w = 100$ | IdentPal $w = 1$ | $H$ |
|---|---|---|---|---|---|
| i. | çéri | $-1$ | | $-1$ | $-1 \times 100 + -1 \times 1 = \mathbf{-101}$ |
| ii. | ☞ xéri | | $-1$ | | $-1 \times 100 = \mathbf{-100}$ |

Restrictiveness in HG requires that the weights of Output-based constraints be high while the weights of Input-Output-based constraints be kept as low as possible; in the initial state, before any learning has occurred, this lowest possible weight for Input-Output-based constraints is zero. As the contrast between (17) and (18) shows, this initial weight of zero for Input-Output-based constraints is necessary to ensure

---

[7] A similar problem exists for error-driven OT learning algorithms that pool the marks of tied constraints (see Jesney and Tessier 2007).

that the productions of the beginning learner's grammar are driven by Output-based considerations.

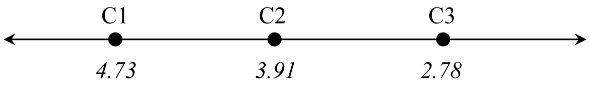
### 3.2 The Gradual Learning Algorithm for HG

Because Harmonic Grammar is a weighted constraint system, it is natural to consider a learning algorithm that will gradually adjust the weights of constraints on the basis of the target language to which the learner is exposed. Here we adopt a version of the Gradual Learning Algorithm (Boersma 1998; Boersma and Hayes 2001; Boersma and Levelt 2003), which initially applied this type of approach to ranked OT grammars, and has been more recently applied to Harmonic Grammars (Boersma and Pater 2007, 2008; Boersma and Weenink 2009).

As Boersma and Levelt (2000, 2003) demonstrate with ranked constraints, the GLA is sensitive to the frequency of different types of input-output pairs and can replicate the stage-like development seen in child language (see also Curtin and Zuraw 2002). Like some other algorithms (notably Error-Driven Constraint Demotion—Tesar and Smolensky 1998, 2000), the GLA responds to each piece of data as it is encountered rather than waiting for all relevant data before altering the grammar. Furthermore, the GLA does not require that mark-data pairs be stored for use in learning over time; only the numerical constraint values are retained (cf. the BCDA of Prince and Tesar 2004, and the LFCD algorithm of Hayes 2004). The combination of Harmonic Grammar and the GLA presented here retains the benefits identified in earlier work, and also converges in certain cases that are problematic with ranked constraints (Boersma and Pater 2008; Pater 2008). We return to explicit comparisons in Sect. 5.

The GLA treats constraints as arrayed along a continuum like that in (20) below:

(20)



In a strict-domination OT system, these numerical values are converted to a ranking on each application of EVAL; in a weighted constraint system, no conversion is necessary. Instead, the numerical values associated with the constraints are simply taken as their weights, and evaluation occurs as discussed in Sect. 2.1.

The assignment of numerical values to constraints allows for gradual adjustment of the constraints' relative importance in the grammar based on evidence from the target language. In the GLA, this adjustment is *error-driven*, with learning triggered whenever there is a mismatch between the forms found in the ambient language and those produced by the learner's current grammar—such as the /çéri/ → *[xéri] mapping of the previous section. Learning occurs only in this type of mismatch scenario; no negative evidence, implicit or explicit, is used.

To illustrate the workings of the GLA we return to the Greek example, where the choice between two outputs in the initial state has resulted in the wrong fricative used before front vowels:

(21) *Initial-state treatment of Greek fricatives: making an error*

| | /çéri/ | *Ç $w = 100$ | *XI $w = 100$ | IDENTPAL $w = 0$ | $H$ |
|---|---|---|---|---|---|
| i. | ✓ çéri | L −1 | | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ xéri | | W −1 | W −1 | $-1 \times 100 + -1 \times 0 = \mathbf{-100}$ |

Selection of a non-target form is an error from which learning can occur. We follow Tesar and Smolensky (1998, 2000); Prince (2002); Becker (2009); and Potts et al. (2010), in labelling the desired target form—[çéri] in (21)—the winner; this label reflects the fact that in the target grammar this form will win in competition against all other candidates. It is marked in tableau with the symbol '✓'. Conversely, the non-target form that the learner produces—[xéri] in (21)—is labelled the loser, since in the target grammar this form will lose out to the winner. It is marked in tableaux with the symbol '☞', consistent with the fact that is the optimal output according to the current grammar. Ws and Ls, indicating whether each constraint 'prefers' the winner or loser, are added to the constraint violations (see Prince 2002 on this notation).

Given an error from which to learn and a set of initial constraint weights, the Harmonic Grammar version of the Gradual Learning Algorithm (the HG-GLA) adjusts these weights until the target form and the form produced by the learner are identical. This is accomplished by increasing the weight of constraints that prefer the winner, and decreasing the weight of constraints that prefer the loser. For the error in (21), this means increasing the weights of *XI and IDENTPALATAL, and decreasing the weight of *Ç.

The precise formulation of the HG-GLA procedure is given in (22)—for discussion, see especially Boersma and Pater (2008).[8]

(22) *HG-GLA learning procedure* [9]

a. The learner is given an input form and a corresponding 'winner' output form, $W$, from the ambient language.
b. The learner takes the input and feeds it to its grammar to determine which output candidate the current grammar deems optimal.
c. If the output candidate selected by the current grammar is identical to the observed output $W$, the grammar remains unchanged.
d. If the output candidate selected by the current grammar is not identical to $W$ (i.e., there is an error), then this output is deemed a 'loser' $L$, and the weight $w_i$ of each constraint $C_i$ is adjusted by:

---

[8] The HG-GLA procedure differs from that applied to ranked constraints—i.e., the OT-GLA (Boersma 1998; Boersma and Hayes 2001; and also Boersma and Pater 2008)—in adjusting the constraint values by an amount that varies based on the *difference* between the violation scores of the winner and the loser ($v_i(W) - v_i(L)$). This is necessary because in HG multiple violations of a single lower-weighted constraint can gang up to overcome a single violation of a higher-weighted constraint, making the number of violations of each constraint directly relevant. As Pater (2008) notes, the HG-GLA update rule is the Perceptron update rule of Rosenblatt (1958). See Boersma and Pater (2008) for simulations and further discussion.

[9] We are grateful to an anonymous reviewer for careful improvement of this procedure.

- subtracting the violation score of the loser, $v_i(L)$ from the violation score of the winner, $v_i(W)$;
- multiplying the result by a plasticity value $r$, where $r > 0$;
- adding the resulting product to the weight of the constraint: $w_i{}^{\text{new}} = w_i{}^{\text{old}} + r(v_i(W) - v_i(L))$.

In our current example the first constraint, *Ç, is violated once by the winner and not at all by the loser; subtracting the violation score of the loser, 0, from the violation score of the winner, $-1$, gives a value of **−1**. The same calculation for *XI and IdentPalatal, each of which are violated once by the loser and not at all by the winner, gives a value of **+1**.

(23) *Grammar with non-target-like output selected*

| | /çéri/ | *Ç<br>$w = 100$ | *XI<br>$w = 100$ | IdentPal<br>$w = 0$ | $H$ |
|---|---|---|---|---|---|
| i. | ✓ çéri (winner) | $-1$ | | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ xéri (loser) | | $-1$ | $-1$ | $-1 \times 100 + -1 \times 0 = \mathbf{-100}$ |
| | $\boldsymbol{v(W) - v(L)}$: | $-1$ | $+1$ | $+1$ | |

Multiplying these values by the plasticity $r$ and adding the product to the initial constraint weights has the result of *decreasing* the value of *Ç and *increasing* the value of *XI and IdentPalatal. This is the desired effect that we began with, and it is the general GLA pattern: given an error, the HG-GLA consistently increases the weights of winner-favoring constraints and decreases the weights of loser-favoring constraints.

Assuming that $r$ is small, a single application of the HG-GLA procedure changes the output of the grammar very little, merely moving the Harmony scores of the winner and loser slightly closer together. If inputs are consistently fed to the procedure in (22) above, however, the accumulation of small adjustments will eventually bring the grammar to the point where no new errors are made. When the target form and the form produced by the learner are identical, constraint weights are no longer adjusted. To simulate this full learning path, one more piece of the story is required: a precise value for $r$. As Sect. 3.3 explains, it is through adjustments to this plasticity value that we perpetuate the initial-state bias discussed in Sect. 3.1.

### 3.3 Biased plasticities in the HG-GLA

The HG-GLA procedure will alter the initial weights of the constraints based on positive evidence from the target language. Ensuring restrictiveness in the final grammar according to the guidelines in (15), however, requires that the weights of Input-Output-based constraints remain as low as possible, while the weights of unviolated Output-based constraints remain high. We propose that this bias is perpetuated through different plasticity values associated with the two types of constraints; in particular, the plasticity of Input-Output-based constraints must be smaller than that of Output-based constraints, causing them to move only slowly from their initial weight of zero. As we will show, this difference allows Input-Output-based constraints to retain relatively low weights even when they are crucially satisfied in the target language, thereby preventing undesirable gang effects that threaten restrictiveness. (For discussion of persistent biases in ranked-constraint learning, see Prince and Tesar 2004; Hayes 2004; Tessier 2007.)

To show the necessity of this persistent bias, we return to the previous English-based example from Sect. 2.2. The target language here typically allows a place contrast between [t] and [k], but strictly disallows clusters composed of a coronal stop followed by a lateral (*TL). With the constraints from Sect. 2, the target language requires the weighting conditions in (24).

(24) a. $w(\text{Ident}) > w(\text{*K})$ … dorsal (and coronal) place is generally permitted.
b. $w(\text{*TL}) > w(\text{Ident}) + w(\text{*K})$ … except before laterals, where coronals are disallowed.

For this target language, the learner's errors in the initial state will come from attempts to repair marked non-coronal place—e.g., /ka/ → [ta]. As illustrated below, these errors are a consequence of the initial weight of the Markedness constraint *K being greater than the initial weight of the conflicting IO-Faithfulness constraint IdentPlace, counter to the target weighting condition (24.a).

(25) *Errors result in w(*K) decreasing and w(IdentPlace) increasing*

| | /ka/ | *TL $w = 100$ | *K $w = 100$ | IdentPlace $w = 0$ | $H$ |
|---|---|---|---|---|---|
| i. | ✓ ka | | L $-1$ → | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ ta | | | ← W $-1$ | $-1 \times 0 = \mathbf{0}$ |

Given the HG-GLA learning procedure, these errors result in the weight of the loser-favoring constraint *K decreasing, indicated with '→', and the weight of the winner-favoring constraint IdentPlace increasing, indicated with '←'. Because no [tl] sequence is ever observed in the target language, *TL is not implicated in any errors and its weight remains unchanged. Continuous adjustments to the weights of *K and IdentPlace will eventually allow the weight of IdentPlace to overcome the weight of *K, reflecting the pattern in (24.a) and causing learning to cease. This leaves open the question, however, of whether this end state will be sufficiently restrictive when faced with input forms that contain illicit /tl/ sequences that were never observed during the course of learning. In other words, will this end state also reflect the pattern in (24.b)?

In fact, the desired restrictive result is only ensured if the weight of IdentPlace remains low enough that it cannot gang up with *K to overcome the weight of the Output-based constraint *TL that is never violated in the target language. Simply assigning different initial weights to Input-Output-based and Output-based constraints is not adequate to ensure this; the low weight of Input-Output-based constraints must be perpetuated. We model this by assigning a lower plasticity value to Input-Output-based constraints than to Output-based constraints, thereby ensuring that the weights of Input-Output-based constraints move away from zero relatively slowly. We will ask shortly exactly *how much* more slowly the weights of Input-Output-based constraints must change in order to ensure restrictiveness, but for now we will assign a plasticity of $r_{\mathrm{IO}} = 0.2$ to Input-Output-based constraints and a plasticity of $r_{\mathrm{O}} = 1.0$ to Output-based constraints. These plasticities mean that every time the HG-GLA learner processes the error in (25), the weight of *K will decrease by 5 times as much as the weight of IDENTPLACE increases. A single error will therefore adjust the grammar as in (26) below.

(26) *Adjusted grammar with biased plasticities after one error*

| | /ka/ | *TL<br>$w = 100$ | *K<br>$w = 99$ | IDENTPLACE<br>$w = 0.2$ | *H* |
|---|---|---|---|---|---|
| i. | ✓ ka | | L −1 → | | $-1 \times 99 = \mathbf{-99}$ |
| ii. | ☞ ta | | | ← W −1 | $-1 \times 0.2 = \mathbf{-0.2}$ |

Repetition of this process will eventually bring the learner to a stable end state with the representative constraint values in (27).[10] No further errors are made at this stage, so no further learning will occur.

(27) *End state with biased plasticities—correct output for observed data*

| | /ka/ | *TL<br>$w = 100$ | IDENTPLACE<br>$w = 16.8$ | *K<br>$w = 16$ | *H* |
|---|---|---|---|---|---|
| i. | ✓☞ ka | | | −1 | $-1 \times 16 = \mathbf{-16}$ |
| ii. | ta | | −1 | | $-1 \times 16.8 = \mathbf{-16.8}$ |

The crucial point now is the treatment of forms like /tla/ that are never encountered during learning. As (28) shows, when such an illicit form is submitted to this grammar it is properly rejected; the end-state grammar is appropriately restrictive.

(28) *End state with biased plasticities—**restrictive***

| | /tla/ | *TL<br>$w = 100$ | IDENTPLACE<br>$w = 16.8$ | *K<br>$w = 16$ | *H* |
|---|---|---|---|---|---|
| i. | ✓☞ kla | | −1 | −1 | $-1 \times 16.8 + -1 \times 16 = \mathbf{-32.8}$ |
| ii. | tla | −1 | | | $-1 \times 100 = \mathbf{-100}$ |

This restrictive result is dependent upon the use of biased plasticities. To see this, we can compare the situation above to the behavior of a learner where Output-based and Input-Output-based constraints have identical plasticities—i.e., $r = 1$ for both classes of constraints; this alternative learner will not be able to reproduce the pattern in (28). For this second learner, the error in (25) still causes the weight of IDENTPLACE to increase until it becomes greater than the weight of *K, but the resulting constraint weights are much higher than they were in the previous case.

---

[10] Comprehensive details of simulations in Praat (Boersma and Weenink 2009) are given in Sect. 3.4.

(29) *End state with $r = 1.0$ for all constraints—still correct output for observed data*

| | /ka/ | *TL<br>$w = 100$ | IDENTPLACE<br>$w = 51$ | *K<br>$w = 49$ | *H* |
|---|---|---|---|---|---|
| i. | ✓☞ ka | | | −1 | −1 × 49 = **−49** |
| ii. | ta | | −1 | | −1 × 51 = −**51** |

As (29) shows, these weights work well for the observed data, but they do not treat unobserved marked structures properly. When fed the novel input /tla/ as in (30), outputs [kla] and *[tla] have equal Harmony scores. The result is a non-restrictive grammar with free variation between [kl] and [tl] onsets. While the grammar learned using the same plasticity value for all constraints selects the correct output for *attested* forms, it is not restrictive enough to rule out faithful mappings for input forms with unattested /tl/ onset sequences.

(30) *End state with $r = 1.0$ for all constraints—**unrestrictive***

| | /tla/ | *TL<br>$w = 100$ | IDENTPLACE<br>$w = 51$ | *K<br>$w = 49$ | *H* |
|---|---|---|---|---|---|
| i. | ✓☞ kla | | −1 | −1 | −1 × 51 + −1 × 49 = −**100** |
| ii. | ☞ tla | −1 | | | −1 × 100 = −**100** |

The general danger illustrated here is that Output-based constraints whose values are unaffected by the learning process (like *TL) may not remain inviolable in the face of gangs of constraints whose weights have been adjusted during the learning process (like IDENTPLACE and *K in (30) above). The crucial element here is the *crossover point* for winner- and loser-favoring constraints whose relative weighting is reversed in the course of learning—here, the point at which the weight of IDENTPLACE exceeds that of *K. Assuming that these constraints are not implicated in any other errors, it is at this point that the weights of these constraints stop changing. If the sum of the weights of IDENTPLACE and *K at the crossover point is *equal to or greater than* the unchanged weight of *TL, the end-state will not be sufficiently restrictive. When $r = 1.0$ for all constraints, the summed weights of IDENTPLACE and *K at the crossover point is 100—the same weight as *TL—which leads to the unrestrictive result seen in (30). If $r_{\mathrm{O}} = 1.0$ for Output-based constraints and $r_{\mathrm{IO}} = 0.2$ for Input-Output-based constraints as in (28), however, the summed weights of IDENTPLACE and *K at the crossover point is 32.8; this is less than the weight of *TL, and the result is appropriately restrictive.

Assigning a lower plasticity value to Input-Output-based constraints than to Output-based constraints systematically lowers the crossover point for conflicting constraints, keeping the weights of Input-Output-based constraints low and reducing the likelihood of undesirable gang effects. Insofar as we have been able to determine, ensuring restrictiveness through this means requires two things. First, the crossover point for any set of constraints must be low enough that even if all of the Input-Output-based constraints were able to gang up with all of the violable Output-based constraints they would not be able to overcome the weight of an inviolable Output-based constraint. It is not possible to set a single crossover threshold, because the crucial comparison is between the weight of unviolated Output-based constraints and the summed weight of all constraints that prefer an intended loser, which cannot be determined *a priori*. The plasticity $r_{\mathrm{IO}}$ of Input-Output-based constraints must therefore be less than the plasticity $r_{\mathrm{O}}$ of Output-based constraints divided by the product of the number of Output-based constraints ($n_{\mathrm{O}}$) and the number of Input-Output-based constraints ($n_{\mathrm{IO}}$).[11]

(31) *Plasticity of Input-Output-based constraints required for restrictiveness*

$$r_{\mathrm{IO}} < \frac{r_{\mathrm{O}}}{n_{\mathrm{IO}} \times n_{\mathrm{O}}}$$

Second, the plasticity $r_{\mathrm{O}}$ of Output-based constraints must be small enough relative to initial constraint weights to ensure that learning is gradual; an Output-based constraint should only attain a weight less than the summed weight of all of the Input-Output-based constraints after multiple learning trials.[12] Given the number of constraints used in our simulations, setting the plasticity to 1 for Output-based constraints and to 0.2 for Input-Output-based constraints ensures restrictiveness in all of the cases discussed here.[13]

### 3.4 Simulations and summary

The previous three sections presented the details of the learner that we use to model the restrictive and gradual acquisition of Harmonic Grammars. Two classes of constraints are distinguished in this model: Output-based constraints and Input-Output-based constraints. Restrictiveness in the end-state grammar requires that the weights of Input-Output-based constraints be kept as low as possible and that the weights of Output-based constraints that are never violated remain high. We implement this requirement in two ways within the framework of the error-driven HG-GLA (Boersma and Pater 2007, 2008; Boersma and Weenink 2009). First, the initial weights of the two classes of constraints are distinguished: Input-Output-based constraints have an initial weight of 0, while Output-based constraints have a high initial weight. Second, the plasticity of Input-Output-based constraints is less than that of Output-based constraints, serving to keep the weights of Input-Output-based constraints relatively low throughout learning.

The learning paths that we report in the rest of this paper are based on HG-GLA simulations conducted in Praat (Boersma and Weenink 2009). We employ the 'LinearOT' decision mode (cf. Keller 2000, 2006), which limits constraint weights to

---

[11] This assumes that constraint violations trade in a strictly one-to-one fashion. The addition of noise introduces further complications that we do not address here.

[12] Thanks to an anonymous reviewer for pointing out the importance of the relationship between the initial constraint weights and the plasticity of Output-based constraints.

[13] An alternative method of implementing this persistent bias, suggested to us by Adam Albright, is exponential decay of constraint weights. On this account, constraint weights would be independently decreased with each application of the learning procedure, on top of any adjustment accorded by the update rule in (22). For independent discussion and related results, see Magri (2007); for a suggestion of how decay might be used to handle outstanding issues relating to later morphophonological learning see Hayes (2004:189) and Tessier (2007:281–283).

values greater than or equal to zero (weights of less than zero are interpreted as zero in this approach). With this decision mode, the HG-GLA procedure in (22) corresponds to Praat's 'Symmetric All' re-ranking strategy. In all of our simulations, plasticity is set to 1.0 for Output-based constraints (Markedness and OO-Faithfulness) and to 0.2 for Input-Output-based constraints (IO-Faithfulness); this plasticity is not altered during the course of learning ('Number of plasticities' = 1). For simplicity, no noise is added to the constraint weights ('Evaluation noise' = 0); as alluded to earlier, however, a noisy version of Harmonic Grammar typically achieves results similar to those reported here (see Jesney and Tessier 2009 for an example). Readers who wish to replicate these simulations, and those in Sect. 3.3, can do so using the input files available at: http://www.ualberta.ca/~annemich/JesneyTessier_GLACollections/.

## 4 Restrictive HG-GLA learning and intermediate stages in language acquisition

The discussion in Sect. 3 focused on our means of implementing a gradual HG learner that arrives at appropriately restrictive end-state grammars based on positive evidence alone. With this learner in hand, we can now also make predictions about intermediate states—stages in a learner's phonological development between the initial and final stages that show the influence of our implementation. The primary goal of this section is to explore these predictions through case studies. We present three examples of developing L1 grammars reported in the literature, and in each case we demonstrate via simulations how our HG-GLA learner passes through the attested stages as a result of its normal workings. These examples highlight the ease with which the learner proposed in Sect. 3 reaches these intermediate stages without biases that distinguish between anything other than Output-based and Input-Output-based constraints. The first case study focuses on the interaction of Specific and General IO-Faithfulness constraints; ensuring restrictiveness when these constraints are involved has been notoriously problematic for previous OT learners (for discussion see Hayes 2004; Prince and Tesar 2004; Tessier 2007). The second and third case studies consider the effects of Output-Output-Faithfulness constraints; our approach achieves results here that are similar to those achieved in other frameworks, but without the need for biases that distinguish between OO-Faithfulness and Markedness constraints. We will return to overt comparisons with other approaches in Sect. 5.

Before introducing the child data used in this and subsequent sections, a methodological and ideological waiver is in order. While the emergent stages discussed in this paper are taken from the reported productions of real children, this paper is primarily a learnability study, and as such it abstracts away from a series of potential confounds. Most importantly, we set aside the possibility that the intermediate stages discussed here may be the result of extragrammatical articulatory or perceptual difficulties rather than the structure of the grammar.[14] Nonetheless, we take these data as a good starting point for the investigation of potential developmental paths in constraint-based error-driven learning and for the testing of predictions made by various models.

### 4.1 Specific *vs.* General IO-Faithfulness and the acquisition of marked structures

We begin by considering the effects of General and Specific IO-Faithfulness constraints in developing grammars, focusing particularly upon what Tessier (2007, 2009) labels Intermediate Faith (IF) stages. IF stages emerge during language acquisition when the target language admits a given marked structure in all contexts, but the child's developing grammar restricts that marked structure to a privileged position. Such stages are well-documented in the child language development literature for a number of languages, with a range of prominent positions being targeted for the preservation of marked structures (see, e.g., Chambless 2006; Goad and Rose 2004; Jesney and Tessier 2009; Kehoe 2000; Marshall et al. 2009; Revithiadou and Tzakosta 2004; Rose 2000; Tessier 2007, 2009). In Harmonic Grammar, IF stages can be modeled as scenarios where the summed weight of the relevant General and Specific IO-Faithfulness constraints is sufficient to overcome Markedness, but the weight of the General IO-Faithfulness constraint alone is not. This is contrasted with the adult target language scenario, where the weight of the General IO-Faithfulness constraint alone is greater than the weight of the conflicting Markedness pressures.

(32) *Weighting conditions of the child's IF stage*

$w(\textsc{Markedness}) > w(\textsc{GeneralFaith})$ … marked structures are disallowed

$w(\textsc{GeneralFaith}) + w(\textsc{SpecificFaith}) > w(\textsc{Markedness})$ … except as faithful mappings in privileged positions.

(33) *Weighting conditions of the target grammar*

$w(\textsc{GeneralFaith}) > w(\textsc{Markedness})$ … marked structures are allowed in all contexts.

As an example of a naturally-emerging IF stage, we consider the production of complex onset clusters by two children learning Québécois French, as reported by Rose (2000). While French allows complex onsets in both stressed and unstressed syllables, Rose (2000) documents an emergent stage for these two children where complex onsets appear only in stressed syllables—a privileged context (Beckman 1998; Curtin 1999, 2001; see also Goad and Rose 2004).

(34) a. *French IF stage*

Stressed complex onsets retained: /CV.ˈ**CCV**/ → [CV.ˈ**CCV**]

Unstressed complex onsets reduced: /**CCV**.ˈCV / → [**CV**.ˈCV], *[**CCV**.ˈCV]

---

[14] At least in the case of Amahl discussed in Sect. 4.2.2, there is strong evidence that articulatory difficulty cannot be responsible for this particular emergent pattern, although perceptual errors may remain a possibility (for discussion, see Dinnsen et al. 2001; Dinnsen and McGarrity 2004; Jesney 2005; Macken 1980).

b. *Data from IF stage: Théo at 2;05.29 - 2;11.29* (Rose 2000:134)

| *Stressed syllables: retained* | | | *Unstressed syllables: reduced* | | |
|---|---|---|---|---|---|
| /′**gʁ**o/ | [′**gʁ**o] | ‘big’ | /**tχ**ak.′tœʁ/ | [**t**a.′tœ$^{ʊ}$] | ‘tractor’ |
| /′**tχ**ɛ̃ | [′**kχ**ɛ] | ‘train’ | /**gʁ**y.′jo/ | [**k**$^{\mathbf{h}}$œ.′jɔ] | ‘oatmeal’ |
| /′**kl**e/ | [′**kx**i] | ‘key’ | /**tχ**u.′ve/ | [**k**ʊ.′βi] | ‘found’ |
| /′**pl**œʁ/ | [′**pl**œ$^{ʊ}$] | ‘s/he cries’ | /′**kχ**ɛm.**gl**a.′se/ | [′**kχ**a$^{i}$.**n**a.′se] | ‘ice cream’ |

Largely following Rose (2000) and Tessier (2007, 2009), we analyze this emergent stage using a Markedness constraint *COMPLEXONSET that militates against consonant-consonant sequences in the syllable onset, and two Faithfulness constraints militating against consonantal deletion—one general (MAX) and one specific to stressed syllables (MAX-$'\sigma$). Given our initial stage assumptions, our learner of French begins with *COMPLEXONSET weighted at 100, and both MAX and MAX-$'\sigma$ constraints weighted at zero. With these weights, a clear pattern of cluster reduction in all environments is predicted in the initial state. (Differences between target and child outputs other than the number of onset segments are ignored here.)

(35) *French onset clusters are all reduced in the initial state*

a. *Stressed syllable*

| | /tχɛ̃/ ‘train’ | *COMPLEX $w = 100$ | MAX $w = 0$ | MAX-$'\sigma$ $w = 0$ | *H* |
|---|---|---|---|---|---|
| i. | ✓ ′kχɛ | L −1 → | | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ ′kɛ | | ← W −1 | ← W −1 | $-1 \times 0 + -1 \times 0 = \mathbf{0}$ |

b. *Unstressed syllable*

| | /tχu.ve/ ‘find’ | *COMPLEX $w = 100$ | MAX $w = 0$ | MAX-$'\sigma$ $w = 0$ | *H* |
|---|---|---|---|---|---|
| i. | ✓ kχʊ.′βi | L −1 → | | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ kʊ.′βi | | ← W −1 | | $-1 \times 0 = \mathbf{0}$ |

Using the HG-GLA learning procedure, these initial-state errors cause the weights of the IO-Faithfulness constraints MAX and MAX-$'\sigma$ to increase and the weight of the Markedness constraint *COMPLEX to decrease. The weight of the general MAX constraint changes more quickly than the weight of the specific MAX-$'\sigma$ constraint, because MAX is implicated in both errors (35.a, b) whereas MAX-$'\sigma$ is implicated in only (35.a).

Learning based on these errors was simulated in Praat using the biased plasticities discussed in Sect. 3.3. With an equal distribution of the two input forms, the first change in the optimal output forms selected occurs after approximately 78 learning trials, with only minimal variation in the constraint values at this stage.[15] A representative example of the weights is shown in (36), along with the effects of these weights in the privileged and non-privileged environments.

(36) *French IF stage—cluster simplification in unstressed syllables only*
a. *Stressed syllable*

| | /tχɛ̃/ | *Complex $w = 22$ | Max $w = 15.6$ | Max-ˈσ $w = 7.6$ | *H* |
|---|---|---|---|---|---|
| i. | ✓☞ ˈkχɛ | $-1$ | | | $-1 \times 22 = \mathbf{-22}$ |
| ii. | ˈkɛ | | $-1$ | $-1$ | $-1 \times 15.6 + -1 \times 7.6 = \mathbf{-23.2}$ |

b. *Unstressed syllable*

| | /tχu.ve/ | *Complex $w = 22$ | Max $w = 15.6$ | Max-ˈσ $w = 7.6$ | *H* |
|---|---|---|---|---|---|
| i. | ✓ kχʊ.ˈβi | L $-1 \rightarrow$ | | | $-1 \times 22 = \mathbf{-22}$ |
| ii. | ☞ kʊ.ˈβi | | $\leftarrow$ W $-1$ | | $-1 \times 15.6 = \mathbf{-15.6}$ |

This emergent pattern is precisely the one displayed by Théo in (34) above. Onset clusters are preserved in stressed syllables but continue to be reduced elsewhere. The crucial relations between the constraint weights that result in this IF pattern are summarized in (37).

(37) *Weighting conditions for Théo's IF stage*
a. $w$(*Complex) > $w$(Max)
i.e., $w$(Markedness) > $w$(GeneralFaithfulness)

b. $w$(Max) + $w$(Max-ˈσ) > $w$(*Complex)
i.e., $w$(GeneralFaithfulness) + $w$(SpecificFaithfulness)
> $w$(Markedness)

Learning continues from this stage as the remaining error in (36.b) triggers further increases to the weight of general Max and further decreases to the weight of *Complex. Max-ˈσ is not implicated in this final error and so its weight remains unchanged. The final stage is reached once the weight of Max independently exceeds that of *Complex, allowing marked onset clusters to be realized even in unstressed syllables. At this point, errors cease and the constraint weights are no longer adjusted. For the case at hand, this occurs when the weight of Max is 16.8 and the weight of *Complex is 16.

---

[15] On each learning trial, Praat randomly selects an input from the user-defined distribution; in all of our simulations, each input is assigned equal probability. Variation in the number of trials required to reach a particular stage and in the precise constraint weights after a set number of trials reflects differences in the learner's sampling from the input distribution in different simulations.

(38) *French onset clusters—the final state*

a. *Stressed syllable*

| | /tχɛ̃/ | Max $w = 16.8$ | *Complex $w = 16$ | Max-′σ $w = 7.6$ | $H$ |
|---|---|---|---|---|---|
| i. | ☞ ′kχɛ | | −1 | | −1 × 16 = **−16** |
| ii. | ′kɛ | −1 | | −1 | −1 × 16.8 + −1 × 7.6 = **−24.4** |

b. *Unstressed syllable*

| | /tχu.ve/ | Max $w = 16.8$ | *Complex $w = 16$ | Max-′σ $w = 7.6$ | $H$ |
|---|---|---|---|---|---|
| i. | ☞ kχʊ.′βi | | −1 | | −1 × 16 = **−16** |
| ii. | kʊ.′βi | −1 | | | −1 × 16.8 = **−16.8** |

The crucial point of the French example is that Harmonic Grammar's additive mode of constraint interaction ensures that IF stages emerge during the course of development despite the lack of evidence for such a stage in the target language. Observing target onset clusters in both stressed and unstressed syllables and adjusting constraint values accordingly is sufficient to produce an emergent intermediate stage where clusters are protected *only* in the privileged position. Even though this stage relies on the Specific Faithfulness constraint Max-′σ, the specific-to-general relationship between Max-′σ and Max does not need to be known in advance or calculated and stored by the HG-GLA learner. As noted at the beginning of this section, IF stages of this type are well-documented in the child language development literature. The HG-GLA's ability to model such stages without additional stipulations is thus a significant advantage to this approach.

This developmental result also reveals the mechanism by which restrictive *end-state* grammars with similar positional restrictions are reached. If a grammar similar to the French IF stage *is* the end-state target, errors and learning cease once weighting conditions like those in (37) are achieved; no further adjustments from this point are required.

By way of example, consider what would occur if, rather than acquiring French, Théo were acquiring a target language that was generally like French but, like Théo's IF stage, admitted consonant clusters only in stressed syllables. Rose (2000:69) cites data provided by Harris (1997:363) suggesting that Southeastern Brazilian Portuguese is indeed such a language: [pra.tu], 'plate', [**pa**.tʃi.ɲu], 'small plate'; [li.**vu**], 'book', [li.**vre**.tu] 'small book'. (See also Goad and Rose 2004.) As in the child French case above, initial-state errors for the learner of this alternative language will prompt an increase in the weights of the winner-favoring constraints Max-′σ and Max, and a decrease in the weight of the loser-favoring constraint *Complex. These adjustments will continue until the summed weight of the General and Specific IO-Faithfulness constraints is greater than the weight of the Markedness constraint alone. At this point, the end-state grammar will have been found. In the present system the weight of Max never individually surpasses that of *Complex (see (36) above). This is a desirable result; if forms containing clusters in unstressed syllables are submitted to the final grammar, they are appropriately rejected (as in (36.b)).

Again, this restrictive result is obtained without the learner needing to know or calculate the specific-to-general relationship between MAX-'$\sigma$ and MAX in advance. Section 5 explicitly compares this result to those obtained in OT models. First, though, we consider additional predictions for emergent stages that are made by the HG-GLA model.

## 4.2 Output-Output Faithfulness and the acquisition of marked sequences

This section turns to emergent stages driven by Output-Output Faithfulness constraints (Benua 1997)—i.e., intermediate stages in acquisition where a markedness-driven process or restriction is circumvented just when failure to do so would cause morphologically-derived forms to differ phonologically from their bases. As with the IF stages discussed in Sect. 4.1, these stages arise despite the fact that the target language provides no evidence for any crucial activity of the OO-Faithfulness constraints in question.

Since the OO-Faithfulness patterns discussed here concern the phonology of morphologically-derived words, we must make some assumptions about when morphological relations are learned by children—i.e., when the notions of base and derived form are incorporated into the developing phonology. It is only once the relationship between a base and its derivative is recognized that OO-Faithfulness constraints can be active in the grammar, because it is only then that the learner knows which output forms to compel faithfulness between. Following previous work, we assume that learners begin by acquiring the purely phonotactic restrictions of their target language in the absence of morphological knowledge (see esp. Hayes 2004; Jarosz Snover 2006; Tessier 2007: Chap. 5), and only later revise their phonology to account for morphologically-sensitive patterns. There are thus two stages of phonological development: first, a 'morphology-free' phonotactic period, where no morphological relations are known and OO-Faithfulness constraints never assign violation marks or choose between candidates, and, later, a morphologically-aware stage where OO-Faithfulness constraints assign violation marks and so contribute to candidate selection. The timing of these stages may well vary across children, but we assume that all learners pass through a phonotactic learning stage before developing morphological awareness.[16]

We begin with a simple but welcome result of the HG-GLA model that treats all Output-based constraints—Markedness and OO-Faithfulness—as belonging to a single constraint class. The pattern of interest is one where the high initial weight of OO-Faithfulness results in an intermediate stage where a target marked structure is admitted just in morphologically-derived contexts. As with the French IF stage discussed in Sect. 4.1, this pattern emerges naturally and without evidence for crucial activity of OO-Faithfulness in the target language.

---

[16] We do not assume, however, that phonotactic learning must be *finished* before morphological awareness emerges; indeed, Amahl's case is one where some morphological relations are discovered and OO-Faithfulness becomes active before all of the target phonotactics are mastered. Similarly, we do not assume that all base-derivative relationships are discovered simultaneously, only that the relationship between a particular derived form and its base must have been recognized by the learner in order for OO-Faithfulness to be active.

The data come from Smith's (1973) famous case study of Amahl's acquisition of English, and in particular Amahl's treatment of pre-lateral coronals. As illustrated in (39) below, Amahl passes through a stage that includes a regular process of pre-lateral velarization where non-strident coronal stops and nasals are realized as velars before target laterals. (All examples are from Smith 1973; page numbers are given in parentheses.)

(39) *Amahl's pre-lateral velarization: applies in all morphologically-simple words*

| | | | | | |
|---|---|---|---|---|---|
| puddle | [pʌ**g**əl] | (240) | sandal | [tæ**ŋg**əl] | (243) |
| butler | [bʌ**k**lə] | (216) | gentle | [dɛ**ŋk**əl] | (227) |

This child-specific process in Amahl's grammar can be understood as the effect of a Markedness constraint that prohibits sequences of non-continuant coronal segments, *TəL, outweighing the IO-Faithfulness constraint demanding that input coronal place be faithfully realized in the output (Dinnsen and McGarrity 2004; Jesney 2005).[17] This is precisely what we expect in the initial state. As phonotactic learning progresses, errors cause the weight of Markedness to decrease and the weight of IO-Faithfulness to increase, as has already begun in (40) below. Assuming the learner initially analyzes all inputs as monomorphemic, violations of OO-IdentCoronal are not assessed at this stage, and its weight remains unchanged.

(40) *Amahl's grammar, partway through phonotactic learning—pre-lateral velarization persists*

| | /pʌ**d**əl/ | OO-IdentCor $w = 100$ | *TəL $w = 40$ | IO-IdentCor $w = 12$ | $H$ |
|---|---|---|---|---|---|
| i. | ✓ pʌ**d**əl | | L −1 → | | −1 × 40 = **−40** |
| ii. | ☞ pʌ**g**əl | | | ← W −1 | −1 × 12 = **−12** |

The innovative OO-Faithfulness aspect of Amahl's pattern is illustrated in (41) below. At the same time as pre-lateral velarization affects simple words like *puddle*, it fails to apply in complex words like those in (41.a)—words where the target coronal segment is part of the base and the target lateral is part of an affix. Evidence that this blocking effect is crucially tied to the segment's realization in the base comes from examples like (41.b); when the target coronal is velarized in the base, this velarization persists in the derived form.

---

[17] *TəL is a more general constraint than the *TL constraint of adult English (see Sect. 2). *TəL is violated by any coronal stop + lateral sequence, regardless of syllabification or the presence of an intervening schwa. Amahl's process of pre-lateral velarization is part of the much-discussed "puzzle-puddle-pickle" chain shift. Target non-strident coronals undergo pre-lateral velarization, yielding mappings like /pʌ**d**əl/ → [pʌ**g**əl] (see (39)); target stridents, however, resist pre-lateral velarization even though they are realized as stops in the output—i.e., /pʌ**z**əl/ → [pʌ**d**əl]. For various approaches to this aspect of the data, see Dinnsen et al. (2001); Dinnsen and McGarrity (2004); Jesney (2005); Macken (1980); Smith (1973).

(41) *Amahl's pre-lateral velarization in complex words*

a. *. . . blocked across base-affix boundary*

| | | | | |
|---|---|---|---|---|
| hard ∼ hardly | [haː**d**] | [haː**dl**iː] | *[haː**gl**i] | (228) |
| soft ∼ softly | [sɔf**t**] | [sɔf**tl**iː] | *[sɔf**kl**iː] | (247) |
| tight ∼ tightly | [tai**t**] | [tai**tl**iː] | *[tai**kl**iː] | (256) |

b. *. . . but not if velarization also occurs in the root*

| | | | | |
|---|---|---|---|---|
| gentle ∼ gently | [dɛŋ**k**əl] | [dɛŋ**kl**iː] | *[dɛn**tl**iː] | (227) |

Having established a set of constraint values like those in (40), this pattern of morphological blocking is predicted to arise once morphologically-complex words are actively produced. As soon as Amahl becomes aware of the affix *-ly* and its role in grouping pairs of words like *hard* and *hardly* into the same paradigm, OO-IdentCoronal can be assessed. The added weight of its violation is sufficient for the OO-Faithfulness pattern to emerge.

(42) *Amahl's OO-Faithfulness stage: (40), with morphologically-complex words*

| | /haːd+liː/<br>**base [haː****d****]** | OO-IdentCor<br>$w = 100$ | *TəL<br>$w = 40$ | IO-IdentCor<br>$w = 12$ | $H$ |
|---|---|---|---|---|---|
| i. | √☞ haːdliː | | −1 | | $-1 \times 40 = \mathbf{-40}$ |
| ii. | haːgliː | −1 | | −1 | $-1 \times 100 + -1 \times 12 = \mathbf{-112}$ |

Since the weight of *TəL decreased during the course of phonotactic learning but the weight of OO-IdentCoronal was unaffected, this grammar levels paradigms towards the faithful (but marked) form in derived words like *hardly*. Phonotactic learning remains incomplete, however, and so *TəL still forces unfaithfulness in simple words like *puddle* where OO-Faithfulness is not assessed. The relevant inequalities for this stage are summarized in (43).

(43) *Inequalities for OO-Faithfulness stage blocking child-specific process*

a. $w$(OO-IdentCoronal) + $w$(IO-IdentCoronal) > $w$(*TəL)
   i.e., $w$(OO-Faithfulness) + $w$(IO-Faithfulness)
   > $w$(Markedness)

b. $w$(*TəL) > $w$(IO-IdentCoronal)
   i.e., $w$(Markedness) > $w$(IO-Faithfulness)

At this stage, Amahl's grammar allows coronal stop + lateral sequences in derived words, but continues to make errors like (40) with simple words. These errors force further decreases in the weight of *TəL and further increases in the weight of IO-IdentCoronal, until the weight of IO-IdentCoronal exceeds that of *TəL, and the final state is reached. This learning path illustrates an important prediction of the current model: OO-Faithfulness stages that override child-specific phonological processes like pre-lateral velarization can only arise if morphological awareness emerges *before* phonotactic learning is complete. Amahl's OO-Faithfulness stage would not occur if morphological awareness arose after the weight of *TəL had decreased and the weight of IO-IdentCoronal had increased to the point that the initial-state inequality was reversed—i.e., to the point where $w$(IO-IdentCoronal) > $w$(*TəL).

Unlike the other emergent stages we discuss, Amahl's OO-Faithfulness stage does not strictly rely on additive interaction. Examination of the intermediate stage tableau in (42) reveals that the weight of OO-IdentCor alone ($w = 100$) is sufficient to overcome the weight of *TəL ($w = 40$). In this case, the gang effect between IO-IdentCor and OO-IdentCor is essentially incidental. Regardless, like the French IF stage, this pattern emerges naturally in the HG-GLA without the need for any biases other than those introduced in Sect. 3. The simple distinction between Output-based constraints and Input-Output-based constraints is adequate.

## 4.3 Output-Output-Faithfulness and the acquisition of alternations

We now turn to a second type of emergent stage that relies on OO-Faithfulness. Whereas Amahl's pattern showed that OO-Faithfulness can allow marked structures to emerge in morphologically-complex words before they are allowed throughout the grammar, the example in this section shows that OO-Faithfulness can also cause the learner to break a correctly-acquired phonotactic pattern once morphological awareness emerges. As with the previous cases, this pattern arises naturally and without the need for any biases other than those introduced in Sect. 3.

The data for this case comes from Kazazis' (1969) report of Marina, a child learning the palatal-velar alternation of Modern Greek (see also Hayes 2004). As described in Sect. 2.2, the target language displays a pattern of complementary distribution, with palatal continuants appearing before front vowels and velar continuants appearing elsewhere; this allophonic pattern trumps any OO-Faithfulness pressures for uniform verbal paradigms in the adult language (see (11)). Adding OO-Faithfulness to the set of Output-based constraints under consideration has no impact on the learner's ability to acquire this pattern of complementary distribution; during initial phonotactic learning the learner does not yet know that any two forms are morphologically related, and so a constraint like OO-IdentPalatal assesses no violation marks. On the basis of errors like the one in (44) below (cf. (21)), the HG-GLA learning procedure increases the weights of *XI and IO-IdentPalatal and decreases the weight of *Ç.

(44) *Palatal allophone before front vowels*

| | /çeri/ | OO-IdentPal $w = 100$ | *XI $w = 100$ | *Ç $w = 100$ | IO-IdentPal $w = 0$ | *H* |
|---|---|---|---|---|---|---|
| i. | ✓ çéri | | | L $-1 \rightarrow$ | | $-1 \times 100 = \mathbf{-100}$ |
| ii. | ☞ xéri | | $\leftarrow$ W $-1$ | | $\leftarrow$ W $-1$ | $-1 \times 100 + -1 \times 0 = \mathbf{-100}$ |

In the absence of morphological awareness, OO-IdentPalatal is vacuously satisfied and so retains its initial weight of 100, with no effect upon the selection of optima. As the hypothetical inputs in (45) show, the phonotactic grammar learned selects the correct allophone regardless of input, in accordance with the principle of Richness of the Base (Prince and Smolensky 1993).

(45) *Final state of pure phonotactic learning: w(OO-IdentPal) = 100, but irrelevant*

a. *Palatal allophone before front vowels . . .*

| | /éxete/ | *XI<br>*w* = 101 | OO-IdentPal<br>*w* = 100 | *Ç<br>*w* = 99 | IO-IdentPal<br>*w* = 0.2 | *H* |
|---|---|---|---|---|---|---|
| i. | √☞ éçete | | | −1 | −1 | −1 × 99 + −1 × 0.2 = **−99.2** |
| ii. | éxete | −1 | | | | −1 × 101 = **−101** |

b. . . . *velar allophone elsewhere*

| | /éço/ | *XI<br>*w* = 101 | OO-IdentPal<br>*w* = 100 | *Ç<br>*w* = 99 | IO-IdentPal<br>*w* = 0.2 | *H* |
|---|---|---|---|---|---|---|
| i. | éço | | | −1 | | −1 × 99 = **−99** |
| ii. | √☞ éxo | | | | −1 | −1 × 0.2 = **−0.2** |

At this stage, the phonotactic pattern of Greek has been mastered, and the learner makes no further errors. Kazazis (1969) reports that Marina reached the stage in (45) at some point before age 4;7, with all of her fricative productions being target-like. For a few weeks beginning at age 4;7, however, Marina began to produce a different pattern, with velar fricatives appearing across the board in some verbal paradigms.

(46) *Marina at 4;7—paradigm leveling*

| | | *Target* | *Child* |
|---|---|---|---|
| 2pl. | 'to have' | [é**ç**ete] | [é**x**ete] |
| 2pl. | 'to leave' | [fév**j**ete] | [fév**ɣ**ete] |

Kazazis (1969) describes this pattern as morphologically-specific under-application of a phonological rule—the rule of Modern Greek that turns underlying velars into palatals before front vowels.

For our HG-GLA learner, the emergence of Marina's pattern in (46) is equated with the onset of morphological awareness. Once the learner acquires enough information about the morphological relationship between the forms of the verbs in question, it becomes possible for OO-IdentPalatal to be assessed, penalizing derived forms that differ in palatal place of articulation with respect to their bases. The resulting violations of OO-IdentPalatal contribute to the Harmony scores of the candidates.

Assessment of OO-IdentPal is dependent upon the identification of a morphological base. This is a non-trivial problem, not least of all because in the current case there is no uninflected member of the paradigm (cf. Benua 1997). While this is an interesting issue for further exploration, for present purposes it is sufficient to assume that the learner selects the unmarked first person singular (velar) form of the stem (see Alderete 1999).[18] As shown in (47), once the velar fricative is chosen as the base

---

[18] A consistent means of selecting a base that remains within the spirit of Benua's (1997) original proposal would be to select the stem allomorph that, context-independently, is most harmonic. In the present case, [éç] violates *Ç, while [éx] violates no Markedness constraints; the velar allomorph would thus be deemed the optimal base form against which OO-IdentPal is assessed. For alternative means of determining bases in inflectional paradigms, see Albright (2005, 2008).

segment against which OO-IDENTPAL is assessed, the grammar in (45) automatically produces Marina's emergent OO-Faithfulness stage.

(47) *Emergent Greek OO-Faithfulness stage at onset of morphological awareness*
a. *Target verb root-final palatals realized as velars under pressure of OO-Faithfulness…*

| | /éx+ete/<br>**base /éx/** | *XI<br>$w = 101$ | OO-IDENTPAL<br>$w = 100$ | *Ç<br>$w = 99$ | IO-IDENTPAL<br>$w = 0.2$ | $H$ |
|---|---|---|---|---|---|---|
| i. | ✓ éçete | | L −1 → | L −1 → | L −1 → | $-1 \times 100 + -1 \times 99$<br>$+ -1 \times 0.2 = -\mathbf{199.2}$ |
| ii. | ☞ éxete | ← W −1 | | | | $-1 \times 101 = -\mathbf{101}$ |

b. … *but target palatals still correctly produced in underived words*

| | /çéri/<br>underived | *XI<br>$w = 101$ | OO-IDENTPAL<br>$w = 100$ | *Ç<br>$w = 99$ | IO-IDENTPAL<br>$w = 0.2$ | $H$ |
|---|---|---|---|---|---|---|
| i. | ✓ ☞ çéri | | | −1 | | $-1 \times 99 = -\mathbf{99}$ |
| ii. | xéri | −1 | | | | $-1 \times 101 = -\mathbf{101}$ |

The General Markedness constraint *Ç is central to this result. At the point when OO-IDENTPAL begins to be assessed, phonotactic learning has already raised the weight of *XI above the weight of *Ç. In underived contexts like (47.b), the $w$(*XI) > $w$(*Ç) inequality holds sway and the target allophonic pattern results. In derived contexts like (47.a), however, OO-IDENTPAL is able to gang up with *Ç and force paradigm leveling. Because the weight of *XI has surpassed that of OO-IDENTPAL during phonotactic learning, the Markedness constraint *Ç is crucial to the equation; the weight of OO-IDENTPAL alone is insufficient to overcome the weight of *XI. Thus, this type of emergent OO-Faithfulness stage, which overrides a target language pattern of complementary distribution, is possible only when there is a General Markedness constraint for OO-Faithfulness to gang up with—i.e., only when there is a constraint that has the same favoring relations as OO-Faithfulness.

(48) *Necessary inequalities for OO-Faithfulness stage overcoming target allophony*
a. $w$(*XI) > $w$(*Ç)
i.e., $w$(SPECIFICMARKEDNESS) > $w$(GENERALMARKEDNESS)
b. $w$(OO-IDENTPALATAL) + $w$(*Ç) > $w$(*XI)
i.e., $w$(OO-FAITHFULNESS) + $w$(GENERALMARKEDNESS)
> $w$(SPECIFICMARKEDNESS)

Progression out of this intermediate stage comes via errors like that in (47.a). These errors demonstrate to the learner that OO-Faithfulness is *not* decisive in the target language and that its weight must be decreased; these errors also serve to increase the difference between the weights of the Markedness constraints *XI and *Ç, although this has no effect on the selection of optima for monomorphemic words. Errors cease once constraint weights like those in (49) are reached.

(49) *Final stage: target palatals correctly realized as palatals in derived contexts*

| | /éx+ete/<br>**base /éx/** | *XI<br>$w = 134$ | OO-IDENTPAL<br>$w = 67$ | *Ç<br>$w = 66$ | IO-IDENTPAL<br>$w = 0$ | $H$ |
|---|---|---|---|---|---|---|
| i. | ✓ ☞ éçete | | $-1$ | $-1$ | $-1$ | $-1 \times 67 + -1 \times 66$<br>$+-1 \times 0 = -\mathbf{133}$ |
| ii. | éxete | $-1$ | | | | $-1 \times 134 = -\mathbf{134}$ |

Like Amahl's pattern from Sect. 4.2, this example shows how the emergence of morphological awareness in concert with the assessment of OO-Faithfulness constraints can bring about innovative intermediate stages for an HG-GLA learner. These patterns emerge without the need for any bias specific to OO-Faithfulness constraints; OO-Faithfulness constraints in this model are treated on par with Markedness constraints. Within the HG-GLA model advanced here, the evidence available from the child's errors alone is sufficient to derive the attested effects. The patterns emerge despite the lack of evidence for OO-Faithfulness activity in the target language.

## 5 Comparisons with emergent stages in OT learning

The HG-GLA model developed in this paper builds on previous work connecting child language acquisition data and models of OT grammar learning. As such, the current proposal retains a number of important attributes of this previous work—including the OT-GLA's ability to model gradual change between stages of development and its sensitivity to input frequency. (For thorough discussion of the OT-GLA and language acquisition, see Boersma and Levelt 2003.)

With respect to restrictiveness, the present HG-GLA model draws on the key observation from the OT literature that different classes of constraints must be biased toward different rankings in order to ensure realistic intermediate stages and restrictiveness in end-state grammars. As motivated for OT by Demuth (1995); Gnanadesikan (2004); Pater (1997); and Smolensky (1996), the HG-GLA model discussed here distinguishes between Markedness and IO-Faithfulness constraints, biasing the first toward high weights, and the second toward the lowest possible weights. As we have shown, the additive mode of constraint interaction in the HG-GLA allows this two-way distinction to be generalized; OO-Faithfulness constraints are treated on par with Markedness constraints, forming the class of Output-based constraints, and Specific IO-Faithfulness constraints are treated on par with General IO-Faithfulness constraints, forming the class of Input-Output-based constraints. Biases in the HG-GLA need only refer to these two constraint classes. In OT, on the other hand, biases must distinguish between four classes of constraints—Markedness, OO-Faithfulness, General IO-Faithfulness, and Specific IO-Faithfulness—in order account for the same data (see esp. Hayes 2004; Prince and Tesar 2004; Tessier 2007). In this section we compare our HG-GLA learner to OT learning approaches discussed in the literature; we focus in particular on the OT-GLA (Boersma 1998; Boersma and Hayes 2001; Boersma and Levelt 2003), as it offers the closest comparison to the present HG-GLA model. As we will see, in some cases our proposal provides an improvement on existing accounts, and in others it performs equally.

### 5.1 General and Specific IO-Faithfulness in OT learning *vs.* HG learning

The need for a Specific IO-Faithfulness ≫ General IO-Faithfulness bias in order to capture attested developmental patterns and end-state grammars has been a concern in the OT learning literature (see esp. Hayes 2004; Tessier 2007; and also Prince and Tesar 2004—discussed further below). As we saw in Sect. 4, the HG-GLA system developed here does not require any such bias; patterns of positional restriction that require a Specific IO-Faithfulness ≫ General IO-Faithfulness ranking in OT can be captured in HG through gang effects involving these two types of constraints. With respect to biases in the HG-GLA, all IO-Faithfulness constraints—both General and Specific—are treated alike. As the following discussion demonstrates, this proves to be a significant advantage.

The importance of biases distinguishing between General and Specific IO-Faithfulness constraints in OT learning is made clear by IF stages like that affecting Théo's onset clusters (see Sect. 4.1). Given a target language like French which requires that the General IO-Faithfulness constraint Max dominate the conflicting Markedness constraint *Complex, Théo's IF stage is not predicted to emerge in a basic OT-GLA system where the only bias is Markedness ≫ IO-Faithfulness. The reason for this is quite simple: since General IO-Faithfulness constraints are implicated in a proper superset of the errors that Specific IO-Faithfulness constraints are implicated in, their values increase more quickly, preventing the restrictive IF stage from ever emerging. From the initial state (50.a) where *Complex dominates Max and Max-$'\sigma$, errors cause the learner to pass directly to the French final state where the General IO-Faithfulness constraint Max dominates *Complex (50.c). Because the Specific IO-Faithfulness constraint is promoted more slowly than the General IO-Faithfulness constraint, the learner never passes through a stage with the Max-$'\sigma$ ≫ *Complex ≫ Max ranking required for the IF stage.

(50) a. *OT-GLA French initial state—Markedness dominates IO-Faithfulness*
*Complex ≫ Max, Max-$'\sigma$
b. *Errors*
/CV.′**CCV**/ → [CV.′**CV**] triggers promotion of Max and Max-$'\sigma$
/**CCV**.′CV / → [**CV**.′CV] triggers promotion of Max
c. *French final state—no IF stage emerges*
Max ≫ *Complex ≫ Max-$'\sigma$

Similar problems emerge for the OT-GLA when restrictive final grammars are considered. To see this, we can consider a target language where onset clusters are only allowed in stressed syllables, i.e., [CV.′**CCV**], *[**CCV**.′CV]. In such a language, all cluster reduction errors will implicate both Max and Max-$'\sigma$, and so the values of the two IO-Faithfulness constraints will be adjusted at the same rate during learning. The result is an unrestrictive final state where Max, as well as Max-$'\sigma$, dominates Markedness.

(51) a. *OT-GLA initial state for target language allowing clusters only in stressed syllables*
*COMPLEX ≫ MAX, MAX-′σ
b. *Errors*
/CV.′**CCV**/ → [CV.′**CV**] triggers promotion of MAX and MAX-′σ
c. *Final state—not restrictive; clusters allowed in all contexts*
MAX, MAX-′σ ≫ *COMPLEX

Based on these two illustrations, it is clear that the OT-GLA learner requires something like a SPECIFIC IO-FAITHFULNESS ≫ GENERAL IO-FAITHFULNESS bias that is imposed throughout learning. This bias must persist even when the distribution of errors prefers that the General IO-Faithfulness constraint be promoted above the Specific IO-Faithfulness constraint (see Hayes and Londe 2006, especially Sect. 6.5, for an example of how the OT-GLA without such a persistent bias fails to learn restrictively).

This issue is not unique to the OT-GLA; similar problems are encountered with other error-driven OT learning algorithms, including Error Driven Constraint Demotion (Tesar and Smolensky 1998, 2000). To address this, Prince and Tesar (2004; see also Hayes 2004's LFCD), propose the Biased Constraint Demotion Algorithm (BCDA), which incorporates persistent biases.[19] There is, however, a larger difficulty with *any* bias that distinguishes between General and Specific IO-Faithfulness constraints as is necessary in OT learning: specific-to-general relationships between Faithfulness constraints can be *language-specific* (Prince and Tesar 2004; see also Hayes 2004; Tessier 2007). Prince and Tesar (2004) illustrate this point with the example of a language where word-initial syllables are always stressed, but non-initial syllables can also bear stress in longer words. In this language Positional Faithfulness constraints referring to stressed syllables (e.g., IDENT-′σ) and to word initial position (e.g., IDENT-INITIAL) in fact stand in a general-to-specific relationship; whenever IDENT-INITIAL applies, so too does IDENT-′σ, but not vice versa. The nature of this relationship, however, can only be determined on a language-specific basis, as it is dependent upon the stress system of the language—and another language with different stress patterns could show the *reverse* relation (see Prince and Tesar 2004 for details).

This inherent difficulty of biases that distinguish between Specific and General IO-Faithfulness constraints has been addressed in two ways for non-GLA learners in the OT literature. The first approach is to adopt different biases, based not on constraint definitions, but rather on constraint activity in observed errors, with the aim of achieving the same effects as a Specific/General IO-Faithfulness bias. This is the method of Prince and Tesar (2004), but their Appendix 3 already provides an example

---

[19] This may often prove adequate for cases where the end state instantiates a positional restriction. Prince and Tesar (2004) are explicit, however, in stating that they do not intend for the BCDA to model intermediate stages in language acquisition. It is thus not surprising that this approach encounters difficulties when extended to such data. In the absence of target language evidence favoring the SPECIFIC IO-FAITH ≫ MARKEDNESS ≫ GENERAL IO-FAITH ranking required for the child's IF stage, such patterns cannot emerge. See Tessier (2007, 2009) for a gradual-learning adaptation of the BCDA approach that accounts for IF stages like Théo's in part by imposing a persistent SPECIFIC IO-FAITHFULNESS ≫ GENERAL IO-FAITHFULNESS bias.

where this does not find the most restrictive grammar possible. The other approach is to commit the learner to a continual search for specific/general relations between each pair of contextually-sensitive IO-Faithfulness constraints. This is pursued in Tessier (2007: Chap. 2, Sect. 6) and in Hayes (2004: fn. 31), but requires considerable repetitive analysis that must be redone throughout the course of acquisition. Both of these techniques crucially rely on stored errors, which are not a component of the GLA; such approaches therefore cannot be extended in any straightforward manner to the OT-GLA.

This entire issue disappears in the HG-GLA system where, due to the reliance on additive interaction, no distinction between General and Specific IO-Faithfulness constraints is necessary. As we have seen, all IO-Faithfulness constraints begin with an initial weight of 0 and their weights are gradually increased on the basis of positive evidence from the target language. As a logical necessity, any error that causes the weight of a Specific IO-Faithfulness constraint to increase will cause the weight of General IO-Faithfulness to increase as well. Restrictive IF stages of the type displayed by Théo result from a Specific IO-Faithfulness constraint and its general counterpart ganging up to protect marked structures just in a specific environment. If the target language enforces such a Positional Faithfulness requirement, learning will cease there; if not, learning will continue until the weight of General IO-Faithfulness alone is greater than the weight of Markedness. In both cases, general/specific effects emerge from target-language evidence and additive interaction. The emergence of restrictive grammars is guaranteed for the HG-GLA learner precisely because a superset-subset relationship holds between the constraints in question, though no grammar-internal knowledge of general/specific relationships is necessary for this to occur. While OT learners face considerable difficulty in addressing this type of data, the HG-GLA learner does so without difficulty; we feel this represents a substantive contribution to this area of the field.

Boersma (2006, 2008) proposes a rather different solution to the problem of positional restrictions in OT learning within a bidirectional framework where comprehension and production grammars are learned in parallel. This approach uses a set of Faithfulness constraints that are *all* context-specific—e.g., where faithfulness to consonant place is mediated by IDENTPLACE-ONSET and IDENTPLACE-CODA constraints. There are no General IO-Faithfulness constraints in this approach. With these crucial assumptions, Boersma (2006, 2008) argues that a stochastic OT-GLA learner can learn grammars where IO-Faithfulness constraints referencing privileged positions (such as onset) dominate those referencing non-privileged positions (such as coda)—no biases within the class of IO-Faithfulness constraints are necessary. To quote: "*if* faithfulness constraints are positional (i.e., are conditioned by syllable position), *and* auditory place cues are *on average* more reliable in onset than in coda, *then* the learner will come to rank IDENTPLACE(onset) above IDENTPLACE(coda), *even if* no such ranking is evident in the parents" (Boersma 2008:21).

This approach is necessarily quite different than our own, and to compare all of its crucial differences would take us rather far afield. We can, however, comment on one important point of divergence—namely, that in adopting constraints like IDENTPLACE-CODA, rather than general IDENTPLACE, the set of languages predicted by factorial typology is increased to include some unattested cases. As

Boersma (2008) notes, if the two Ident constraints are held in a fixed ranking of IDENTPLACE-ONSET ≫ IDENTPLACE-CODA, the patterns predicted are the same as those predicted in a system with specific IDENTPLACE-ONSET and general IDENTPLACE. Once this fixed ranking is given up, however, the unattested pattern of place preservation only in coda position becomes available in principle. It is left to the relative strength of the cues in different positions to block such a grammar from emerging. Independent of whether this approach can guarantee that the learner will never choose a pathological ranking, it is clear that this model gives up the goal of a factorial typology that fully matches attested cross-linguistic patterns—a move that may well have unforeseen effects (though cf. Alderete 2008). The HG-GLA approach does not need to adopt constraints like IDENTPLACE-CODA in order to capture positional effects, so it cannot produce the coda-preferential grammar even in principle. Additive constraint interaction, along with a constraint set including General IO-Faithfulness constraints and Specific IO-Faithfulness constraints relativized to privileged positions, is adequate to ensure restrictive intermediate stages and end-state grammars.

## 5.2 OO-Faithfulness in OT learning vs. HG learning

The discussion of Specific and General IO-Faithfulness constraints in the previous section demonstrated that the HG-GLA learner offers a straightforward solution to important issues in constraint-based grammar learning. The HG-GLA learner does not resolve any crucial outstanding issues with respect to OO-Faithfulness constraints; it does, however, attain the same level of data coverage as OT learning approaches in the literature, and it does this without the need for any bias distinguishing between OO-Faithfulness and Markedness constraints.

On the basis of acquisition data like Marina's, it has been argued that a restrictive OT learner must begin with OO-Faithfulness constraints undominated, ranked above even Markedness constraints (Hayes 2004; Tessier 2006, 2007; see also McCarthy's 1998 Richness of the Base argument for an OO-FAITHFULNESS ≫ IO-FAITHFULNESS bias). As compared to the Specific IO-Faithfulness vs. General IO-Faithfulness bias discussed in the previous section, this bias is reasonably simple to implement, given that the class of OO-Faithfulness constraints can be identified in a straightforward fashion and is not subject to language-particular vagaries. Nonetheless, implementation of this bias is not necessarily easy for the OT-GLA learner, as it must be enforced in a persistent fashion throughout learning. Visible effects of OO-Faithfulness in OT often require that OO-Faithfulness outrank Markedness, and that Markedness outrank IO-Faithfulness. In a case like that of Marina, discussed in Sect. 4.3, the necessary OT ranking is OO-FAITH ≫ *XI ≫ *Ç ≫ IO-FAITH. As that section showed, however, during the course of GLA-based phonotactic learning the value of the Markedness constraint *XI increases even as the value of OO-Faith remains unchanged (see (46)). Under HG evaluation Marina's OO-Faithfulness stage emerges regardless, due to the additive interaction of OO-FAITH and *Ç. An OT learner, on the other hand, requires a strict ranking of OO-FAITH ≫ *XI to create Marina's pattern. To ensure this, the OT-GLA learner will therefore require a persistent bias, in which OO-Faithfulness constraints are moved up in tandem with Markedness constraints even in the absence of positive evidence.

Our HG-GLA learner requires no bias distinguishing between OO-Faithfulness and Markedness constraints. The restrictive developmental stages discussed in Sects. 4.2 and 4.3 emerge either from additive interaction, as OO-Faithfulness gangs up with additional constraints that favor the leveled paradigm, or from the simple fact that the weights of conflicting constraints have decreased even as the initial weight of OO-Faithfulness remains unchanged. The HG-GLA learner is thus able to achieve empirical coverage comparable to existing OT approaches while relying on fewer classes of constraints in its biases—a simplification of the overall system.

## 6 Conclusion

This paper has demonstrated that restrictive phonological learning can be modeled with biases that distinguish between only two classes of constraints—Output-based constraints and Input-Output-based constraints—when the mode of constraint interaction is additive as in Harmonic Grammar and learning is gradual as in the GLA. Here these biases were implemented through differences in the initial weights of the two classes of constraints, and through differences in the rates at which the weights of the two classes of constraints are adjusted.

(52) *Two classes of constraints distinguished for restrictiveness in the HG-GLA*
  a. **Output-based** constraints should be weighted high—high enough to remain inviolable if necessary.
  b. **Input-Output-based** constraints should be weighted as low as possible.

The HG-GLA approach presented here both allows restrictive end-state grammars to be learned on the basis of positive evidence, and predicts the natural emergence of a range of attested intermediate stages. These effects are largely the result of additive constraint interaction. Emergent IF stages, where a marked structure is realized faithfully just in a privileged context, are the result of General and Specific IO-Faithfulness constraints ganging up to overcome the weight of a conflicting Markedness constraint. End-state grammars with similar positional restrictions are accounted for through the same type of weighting conditions. Importantly, this approach avoids the need to compute language-particular general/specific relationships.

Emergent effects involving OO-Faithfulness are more varied in nature, though they too can rely on additive interaction. Patterns like Amahl's OO-Faithfulness stage, where a child-specific process is blocked just in derived words, are a consequence of the persistent high weight assigned to OO-Faithfulness. Even as the weights of other constraints change during the course of learning, the high initial weight of OO-Faithfulness remains unaltered in the absence of evidence to the contrary—as with all other unviolated Output-based constraints. Patterns like Marina's OO-Faithfulness stage, where a target language process is blocked in the child's language, rely on this high weight of OO-Faithfulness as well, but also depend upon OO-Faithfulness ganging up with a general Markedness constraint.

That an HG-GLA system need only distinguish between Output-based and Input-Output-based constraints is a significant advantage of this approach. Distinguishing between fewer classes of constraints allows the learner to make fewer *a priori* decisions about the nature of the language being learned—e.g., which constraints are general vs. specific. Furthermore, fewer constraint classes means that fewer initial and persistent differences in constraint weights/rankings need be implemented. At the same time, the HG-GLA model retains the advantages of the OT-GLA, with learning being gradual, error-driven, and sensitive to the nature of the target language data.

There remains much scope for further research into the HG-GLA model. As one example, different means of persistently implementing high weights for Output-based constraints *vs.* lowest possible weights for Input-Output-based constraints should be considered and compared with the plasticity-based approach adopted here. The range and nature of emergent restrictiveness effects in acquisition is similarly in need of further exploration, with many empirical questions remaining to be answered. Nonetheless, it is clear that a combination of additive constraint interaction with a gradual on-line learning algorithm offers considerable advantages, providing novel support for the Harmonic Grammar model and for constraint weighting more generally.

## References

ROA = Rutgers Optimality Archive. http://roa.rutgers.edu.

Albright, Adam. 2005. The morphological basis of paradigm leveling. In *Paradigms in phonological theory*, eds. Laura Downing, Tracy Alan Hall, and Renate Raffelsiefen. Oxford: Oxford University Press.

Albright, Adam. 2008. Inflectional paradigms have bases too: arguments from Yiddish. In *The bases of inflectional identity*, eds. Asaf Bachrach and Andrew Nevins. Oxford: Oxford University Press.

Alderete, John. 1999. Morphologically-governed accent in optimality theory. Amherst, MA: UMass Amherst dissertation [ROA-309].

Alderete, John. 2008. Using learnability as a filter on computable functions: a new approach to Anderson and Browne's generalization. *Lingua* 118: 1177–1220.

Becker, Michael. 2009. Phonological trends in the lexicon: the role of constraints. Amherst, MA: UMass Amherst dissertation.

Beckman, Jill. 1998. Positional faithfulness. Amherst, MA: UMass Amherst dissertation.

Benua, Laura. 1997. Transderivational identity: phonological relations between words. Amherst, MA: UMass Amherst dissertation.

Boersma, Paul. 1998. Functional phonology: formalizing the interactions between articulatory and perceptual drives. University of Amsterdam dissertation.

Boersma, Paul. 2006. The acquisition and evolution of faithfulness rankings. Paper presented at Manchester Phonology Meeting 14, May 2006.

Boersma, Paul. 2008. Emergent ranking of faithfulness explains markedness and licensing by cue. Manuscript, University of Amsterdam [ROA-954].

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1): 45–86 [ROA-348]

Boersma, Paul, and Claartje Levelt. 2000. Gradual constraint-ranking learning algorithm predicts acquisition order. In *Proceedings of the 30th child language research forum*, ed. Eve V. Clark, 229–237. Stanford: CSLI [ROA-361].

Boersma, Paul, and Claartje Levelt. 2003. Optimality Theory and phonological acquisition. *Annual Review of Language Acquisition* 3: 1–50.

Boersma, Paul, and Joe Pater. 2007. Gradual learning of phonotactic constraints. Manuscript, University of Amsterdam & UMass Amherst. Paper presented at 38th meeting of the North East Linguistic Society, October, 2007.

Boersma, Paul, and Joe Pater. 2008. Convergence properties of a gradual learning algorithm for Harmonic Grammar. Manuscript, University of Amsterdam & University of Massachusetts Amherst [ROA-970].

Boersma, Paul, and David Weenink. 2009. Praat: doing phonetics by computer. www.praat.org.

Chambless, Della. 2006. Asymmetries in consonant cluster acquisition. Amherst, MA: UMass Amherst dissertation.

Curtin, Suzanne. 1999. Positional prominence in the acquisition of prosodic structure. In *Proceedings of the annual meeting of the Canadian Linguistics Society*, 89–100.

Curtin, Suzanne. 2001. Representational richness in phonological development. Los Angeles, CA: University of Southern California dissertation.

Curtin, Suzanne and Kie Zuraw. 2002. Explaining constraint demotion in a developing system. In *BUCLD 26: Proceedings of the 26th annual Boston University conference on language development*, eds. Anna H.-J. Do, Laura Domínguez, and Aimee Johansen. Somerville: Cascadilla Press.

Demuth, Katherine. 1995. Markedness and the development of prosodic structure. In *Proceedings of the 25th meeting of the North East Linguistic Society*, ed. Jill Beckman. Amherst: Graduate Linguistic Student Association [ROA-50].

Dinnsen, Daniel A., and Laura W. McGarrity. 2004. On the nature of alternations in phonological acquisition. *Studies in Phonetics, Phonology and Morphology* 10: 23–41.

Dinnsen, Daniel A., Kathleen M. O'Connor, and Judith A. Gierut. 2001. The puzzle-puddle-pickle problem and the Duke-of-York gambit in acquisition. *Journal of Linguistics* 37: 503–525.

Gnanadesikan, Amalia. 2004. Markedness and faithfulness in child phonology. In *Fixing Priorities: Constraints in Phonological Acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 73–108. Cambridge: Cambridge University Press [ROA-67].

Goad, Heather, and Yvan Rose. 2004. Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic. In *Fixing priorities: constraints in phonological acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 109–157. Cambridge: Cambridge University Press [ROA-448].

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT constraint rankings using a Maximum Entropy model. In *Proceedings of the workshop on variation within Optimality Theory*, 111–120. Stockholm: Stockholm University.

Hale, Mark, and Charles Reiss. 1998. Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29(4): 656–683.

Harris, John. 1997. Licensing inheritance: an integrated theory of neutralisation. *Phonology* 14: 315–370.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the early stages. In *Fixing priorities: constraints in phonological acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 158–203. Cambridge: Cambridge University Press [ROA-327].

Hayes, Bruce, and Zsuzsa Londe. 2006. Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology* 23: 59–104.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440 [ROA-858].

Jäger, Gerhard. 2007. Maximum entropy models and stochastic Optimality Theory. In *Architectures, rules, and preferences: a festschrift for Joan Bresnan*, eds. Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson, and Annie Zaenen, 467–479. Stanford: CSLI.

Jarosz Snover, Gaja. 2006. Rich lexicons and restrictive grammars—maximum likelihood learning in Optimality Theory. Baltimore, MD: Johns Hopkins University dissertation [ROA-884].

Jesney, Karen. 2005. Chain shift in phonological acquisition. Calgary, AB: University of Calgary MA Thesis. http://people.umass.edu/kjesney/Jesney2005MA.pdf.

Jesney, Karen, and Anne-Michelle Tessier. 2007. Restrictiveness in gradual learning of Harmonic Grammar. Manuscript UMass Amherst & University of Alberta. Paper presented at the North East computational phonology circle (NecPhon 1), November, 2007.

Jesney, Karen, and Anne-Michelle Tessier. 2009. Gradual learning and faithfulness: consequences of ranked *vs.* weighted constraints. In *Proceedings of the North East Linguistic Society 38*, eds. Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, 375–388. Amherst: GLSA.

Kazazis, Kostas. 1969. Possible evidence for (near-)underlying forms in the speech of a child. In *Papers from the fifth regional meeting of the Chicago Linguistic Society (CLS 5)*, eds. Robert I. Binnick, Alice Davison, Georgia McGreen, and Jerry L. Morgan, 382–388. Chicago: Chicago Linguistic Society.

Kazazis, Kostas. 1972. Some remarks on palatalization in Modern Greek. *Research on Language & Social Interaction* 5(2): 330–332.

Kehoe, Margaret. 2000. Truncation without shape constraints: the latter stages of prosodic acquisition. *Language Acquisition* 8(1): 23–67.

Kehoe, Margaret, and Carol Stoel-Gammon. 1997. Truncation patterns in English-speaking children's word productions. *Journal of Speech, Language, and Hearing Research* 40: 526–541.

Keller, Frank. 2000. Gradience in grammar: experimental and computational aspects of degrees of grammaticality. Edinburgh: University of Edinburgh dissertation.

Keller, Frank. 2006. Linear Optimality Theory as a model of gradience in grammar. In *Gradience in grammar: generative perspectives*, eds. Gisbert Fanselow, Caroline Féry, Ralph Vogel, and Matthias Schlesewsky. Oxford: Oxford University Press.

Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic Grammar—a formal multi-level connectionist theory of linguistic wellformedness: an application. In *Proceedings of the twelfth annual conference of the Cognitive Science Society*, 884–891. Cambridge: Lawrence Erlbaum.

Macken, Marlys A. 1980. The child's lexical representation: the 'puzzle-puddle-pickle' evidence. *Journal of Linguistics* 16: 1–17.

Magri, Giorgio. 2007. The subset problem in Optimality Theory. Manuscript, MIT.

Marshall, Chloe R., and Heather, K.J. van der Lely. 2009. Effects of word position and stress and onset cluster production: evidence from typical development, specific language impairment, and dyslexia. *Language* 85: 39–57.

McCarthy, John J. 1998. Morpheme structure constraints and paradigm occultation. In *Papers from the thirty-second regional meeting of the Chicago Linguistic Society (CLS 32), part 2: the panels*, eds. M. Catherine Gruber, Derrick Higgins, Kenneth Olsen, and Tamra Wysocki, 123–150. Chicago: Chicago Linguistic Society.

McCarthy, John J. 2002. *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.

McCarthy, John J., and Alan Prince. 1995. Faithfulness and reduplicative identity. In *Papers in optimality theory: university of Massachusetts occasional papers*, Vol. 18, 249–384. Amherst: GLSA [ROA-60].

Moreton, Elliott. 2000. Phonological grammar in speech perception. Amherst, MA: UMass Amherst dissertation.

Moreton, Elliott. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84: 55–71.

Pater, Joe. 1997. Minimal violation in phonological development. *Language Acquisition* 6(3): 201–253.

Pater, Joe. 2004. Bridging the gap between perception and production with minimally violable constraints. In *Fixing priorities: constraints in phonological acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 219–244. Cambridge: Cambridge University Press.

Pater, Joe. 2008. Gradual learning and convergence. *Linguistic Inquiry* 39(2): 334–345.

Pater, Joe. 2009. Weighed constraints in generative linguistics. *Cognitive Science* 33: 999–1035.

Pater, Joe, Chris Potts, and Rajesh Bhatt. 2007. Linguistic Optimization. Manuscript, University of Massachusetts Amherst [ROA-872].

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt, and Michael Becker. 2010. Harmonic Grammar with linear programming: from linear systems to linguistic typology. *Phonology* 27(1): 77–117 [ROA-984].

Prince, Alan. 2002. Arguing optimality. In *Papers in optimality theory II*, eds. Andries Coetzee, Angela Carpenter, and Paul de Lacy, Amherst: GLSA.

Prince, Alan. 2003. Anything goes. In *New century of phonology and phonological theory*, eds. Takeru Honma, Masao Okazaki, Toshiyuki Tabata, and Shinichi Tanaka, 66–90. Tokyo: Kaitakusha [ROA-536].

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar. Technical Report, Rutgers University and University of Colorado at Boulder, 1993. Revised version published by Blackwell, 2004 [ROA-537].

Prince, Alan, and Bruce Tesar. 2004. Learning phonotactic distributions. In *Fixing priorities: constraints in phonological acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 245–291. Cambridge: Cambridge University Press [ROA-353].

Revithiadou, Anthi, and Marina Tzakosta. 2004. Markedness hierarchies *vs.* positional faithfulness and the role of multiple grammars in the acquisition of Greek. In *Proceedings of generative approaches to language acquisition (GALA)*. Utrecht: LOT Occasional Series.

Rose, Yvan. 2000. Headedness and prosodic licensing in the L1 acquisition of phonology. Montreal, PQ: McGill University dissertation.

Rosenblatt, Frank. 1958. The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* 65: 386–408.

Smith, Jennifer L. 2000. Positional faithfulness and learnability in Optimality Theory. In *Proceedings of ESCOL 99*, eds. Rebecca Daly and Anastasia Riehl, 203–214. Ithaca: CLC.

Smith, Neilson V. 1973. *The acquisition of phonology: a case study*. Cambridge: Cambridge University Press.

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

Smolensky, Paul, and Géraldine Legendre. 2006. *The harmonic mind: from neural computation to optimality-theoretic grammar*. Cambridge: MIT Press.

Stampe, David. 1973. A dissertation in natural phonology. Chicago, IL: University of Chicago dissertation.

Tesar, Bruce. 1997. An iterative strategy for learning metrical stress in Optimality Theory. In *Proceedings of the twenty-first annual Boston University conference on language acquisition*, eds. Elizabeth Hughes, Mary Hughes, and Annabel Greenhill, 615–626. Somerville: Cascadilla.

Tesar, Bruce, and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29(2): 229–268.

Tesar, Bruce, and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge: MIT Press.

Tessier, Anne-Michelle. 2006. Testing for OO-Faithfulness in artificial phonological acquisition. In *Proceedings of the 30th annual Boston University conference on language development*, eds. David Bamman, Tatiana Magnitskaia, and Colleen Zaller, 619–639. Somerville: Cascadilla [ROA-817].

Tessier, Anne-Michelle. 2007. Biases and stages in phonological acquisition. Amherst, MA: University of Massachusetts Amherst dissertation [ROA-883].

Tessier, Anne-Michelle. 2009. Frequency of violation and constraint-based phonological learning. *Lingua* 119(1): 6–38.

Velleman, Shelley L. 1988. The role of perception in later phonological development. *Applied Psycholinguistics* 9: 221–236.

Werker, Janet, and Richard Tees. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development* 7: 49–63.

Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* 30: 945–982.