Gradual Learning and Faithfulness: Consequences of Ranked vs. Weighted Constraints*

Karen Jesney[1] & Anne-Michelle Tessier[2]

[1]University of Massachusetts Amherst & [2]University of Alberta

1. Overview

This paper investigates a class of restrictive intermediate stages that emerge during L1 phonological acquisition, and argues that these stages are naturally accounted for within a gradual learning model that uses weighted constraints. The particular type of pattern of interest here – Intermediate Faithfulness (IF) stages – involves the preservation of marked structures just in privileged environments. We illustrate this with data from Bat-El (2007), which shows the innovation of morphologically-sensitive phonology during the acquisition of Hebrew. The resulting IF stage displays greater faithfulness in nouns – a privileged context (Smith 2000, 2001) – than in non-nouns. Like other IF stages, it emerges without any direct support from the target grammar (Revthiadou & Tzakosta, Rose 2000, Tessier 2007a).

Our IF-stage analysis makes use of a gradual on-line learner and a grammar of weighted constraints (as in Harmonic Grammar (HG); Legendre et al. 1990, Smolensky & Legendre 2006). In this model, the harmony score H of a candidate R is calculated by multiplying its violations of each constraint $\{C_1(R), C_2(R), \ldots, C_n(R)\}$ by the weights associated with those constraints $\{w_1, w_2, \ldots, w_n\}$ and summing, as in (1). The candidate with the highest resulting harmony score is deemed optimal.

(1) $H(R) = (C_1(R)*w_1) + (C_2(R)*w_2) + \ldots + (C_n(R)*w_n)$

In many situations the HG optimization procedure gives much the same result as the ranked constraints of Optimality Theory (Prince & Smolensky 2004). This is illustrated schematically in (2a), where each candidate violates only one constraint. The input here is the marked element /A/, which occurs in a non-privileged environment in

* Thanks to Joe Pater, Adam Albright, Paul Boersma, Gaja Jarosz, John McCarthy, Johanne Paradis, Pat Pratt, Matt Wolf, and the audiences at NELS38 and NecPhon1. Special thanks to Michael Becker for pointing the authors to the data, and to Avivit Ben-David and Galit Adam.

To appear in *Proceedings of the 38th Meeting of the North East Linguistic Society (NELS 38)*, eds. Muhammad Abdurrahman, Anisa Schardl & Martin Walkow. Amherst, MA: GLSA. [Feb. 29, 2008]

this example. The faithful candidate [A] violates Markedness (*w*=70), giving it a harmony score of -1(70) = -70, while the unfaithful candidate [B] violates General Faithfulness (*w*=50), giving it a harmony score of -1(50) = -50. The harmony score of output [B] is thus higher than that of output [A], and so the marked input structure /A/ maps unfaithfully to [B].[1] This is the same output as would be selected with an OT ranking of Markedness >> General Faithfulness >> Specific Faithfulness.

(2) a. Marked structures **are not** generally permitted $/A_{\text{non-privileged}}/ \rightarrow$ [B]

$/A_{\text{non-privileged}}/$	Markedness $w = 70$	GenFaith $w = 50$	SpecFaith $w = 40$	*H*
A (faithful)	-1			-1(70) = **-70**
☞ B (unfaithful)		-1		-1(50) = **-50**

The analysis of IF stages developed here builds on a property of HG that is *not* shared with OT – the ability multiple lower-weighted constraints to *gang up* to overcome a higher-weighted constraint. In (2b), for example, input /A/ is found in a privileged environment, making the Specific Faithfulness constraint relevant. The faithful candidate [A] violates Markedness, giving it a harmony score of -1(70) = -70, while the unfaithful candidate [B] violates both General Faithfulness and Specific Faithfulness, giving it a harmony score of -1(50) + -1(40) = -90. The harmony score of [A] is thus greater than that of [B], making the faithful candidate [A] the winner in the privileged context.

(2) b. Marked structures **are** permitted in privileged contexts $/A_{\text{privileged}}/ \rightarrow$ [A]

$/A_{\text{privileged}}/$	Markedness $w = 70$	GenFaith $w = 50$	SpecFaith $w = 40$	*H*
☞ A (faithful)	-1			-1(70) = **-70**
B (unfaithful)		-1	-1	-1(50) + -1(40) = **-90**

The constraint weights are the same in (2a) and (2b); it is the difference in constraint violations and the availability of gang effects that gives rise to the IF profile where the marked structure /A/ is mapped faithfully just in privileged environments. This set of mappings can, of course, be generated in OT through a ranking of SpecFaith >> Markedness >> GenFaith. A parallel set of weighting conditions – i.e., *w*SpecFaith > *w*Markedness > *w*GenFaith – will also achieve this effect. The HG optimization procedure does not add to the typology here (cf. Pater, Potts & Bhatt 2007); instead, it simply provides an additional means by which the IF pattern can be generated.

The gradual error-driven HG learner discussed in this paper – i.e., the HG-GLA – naturally passes through the gang-effect-based IF stage described above. The learner begins with a basic *w*Markedness > *w*Faithfulness bias and gradually adjusts the weights of the constraints on the basis of target-language evidence. Under this approach, as the weights of the General and Specific Faithfulness constraints increase, the gang effect IF

[1] Constraint violations are represented as negative numbers (-1 in these tableaux) and constraint weights are limited to non-negative numbers (Keller 2006, cf. Prince 2002, Pater, Potts & Bhatt 2007). This has the effect of making zero the maximum possible harmony score for any candidate.

pattern in (2a) and (2b) naturally emerges and the learner passes easily through the restrictive subset grammar on the way to the target. The same HG-GLA procedure also achieves considerable success in ensuring restrictiveness for end-state grammars (see Jesney & Tessier 2007, Magri 2007, Pater, Jesney & Tessier 2007, Tessier 2007b).

The rest of this paper is organized as follows: Section 2 presents word truncation data from Hebrew acquisition and reanalyzes the morphologically-sensitive stage identified by Bat-El (2007) as an IF stage where nouns are the privileged context. Section 3 spells out how IF stages emerge in the HG-GLA approach, and discusses the difficulties faced by a gradual learner that instead uses ranked OT constraints. Section 4 offers concluding remarks and directions for further research.

2. Stages of Word Truncation in Hebrew Acquisition

Word truncation is a very common type of intermediate stage in L1 phonological development in which the child deletes entire syllables from a target word to create a prosodic shape that is unmarked with respect to stress placement, foot alignment or other factors (see e.g., Demuth & Fee 1995, Fikkert 1994, Pater 1997). Longitudinal studies by Adam (2002) and Ben-David (2001) report two early stages of word truncation in Hebrew acquisition – a Trochaic Stage where words are truncated to a single trochee, followed by a Disyllabic Stage where words are truncated to a single disyllable regardless of stress. This section analyzes these two stages within the HG framework laid out above, and then turns to the grammar of a child discussed by Bat-El (2007) where these two patterns co-exist. We argue that the innovative pattern of this final child represents an IF stage where General Faithfulness and Noun-Specific Faithfulness gang up to overcome conflicting markedness pressures.

2.1 The Hebrew Trochaic Stage

During the Trochaic Stage in Hebrew acquisition all target words are truncated to a single moraic or syllabic trochee built around the input stressed syllable.

(3) The Hebrew Trochaic stage:
/…Sw/ → [(Sw)] /…wS/ → [(S)]

Representative data are given below; mean age of the children at this stage is 1;2 (Avivit Ben-David, p.c.).

(4) Data from the Trochaic Stage (Adam 2002: 72-73)

input: /…Sw/		
target	*child*	*gloss*
ˈse.fer	ˈfe.fe	‘book’
ʃar.ˈʃe.ret	ˈe.tet	‘necklace’
la.ˈre.det	ˈde.det	‘to get down’
lif.ˈto.ax	ˈgo.ax	‘to open’

input: /…wS/		
target	*child*	*gloss*
ka.ˈdur	ˈdur	‘ball’
sa.ˈbon	ˈbon	‘soap’
lis.ˈgor	ˈgor	‘to close’
ne.ʃɪ.ˈka	ˈka	‘kiss’ (n.)

In our analysis two markedness constraints drive this pattern: ALIGN-FT-L, which requires that the left edge of every foot be aligned with the left edge of a Prosodic Word (McCarthy & Prince 1993 *et seq*), and TROCHEE, which requires that feet be left-headed.[2] As illustrated in (5) below, when both of these constraints are weighted more heavily than MAXV, which militates against vowel and therefore syllable deletion, all pre-tonic syllables are elided, yielding the Trochaic Stage.

(5) a. Longer words with penultimate stress are truncated: /wSw/ → [(Sw)][3]

/ʃar.ˈʃe.ret/ 'necklace'	TROCHEE *w*=100	ALIGN-FT-L *w*=100	MAXV *w*=0	*H*
☞ (ˈe.tet)			-1	-1(0) = **0**
ar.(ˈe.tet)		-1		-1(100) = **-100**

(5) b. Disyllables with final stress are also truncated: /wS/ → [(S)]

/ka.ˈdur/ 'ball'	TROCHEE *w*=100	ALIGN-FT-L *w*=100	MAXV *w*=0	*H*
☞ (ˈdur)			-1	-1(0) = **0**
ka.(ˈdur)		-1		-1(100) = **-100**
(ka.ˈdur)	-1			-1(100) = **-100**

The early Trochaic Stage thus requires that the two Markedness constraints each have a greater weight than the conflicting IO-Faithfulness constraint – the familiar M >> F initial state discussed in the OT literature (e.g., Gnanadesikan 2004, Smolensky 1996; see Jesney & Tessier 2007 for discussion of this bias in HG). With MAXV low weighted, syllables are freely truncated to ensure that the ideal trochaic shape is met.

(6) Trochaic Stage: *w*ALIGN-FT-L, *w*TROCHEE > *w*MAXV

2.2 The Hebrew Disyllabic Stage

During the second stage of Hebrew word truncation – the Disyllabic Stage – all words surface with up to two syllables and either initial or final stress. Both trochees and iambs are now permitted, depending upon the location of input stress.

(7) The Hebrew Disyllabic stage:
/…Sw/ → [(Sw)]   /…wS/ → [(wS)]

Representative data from the same children is given in (8); mean age at this stage is 1;4 (Avivit Ben-David, p.c.).

---

2 Adam (2002) uses a somewhat different set of constraints to analyze this data (see Tessier 2007b).

3 To ensure that stress is not shifted to minimize truncation, a faithfulness constraint like IDENT-STRESS must be weighted more highly than MAXV. This stage also requires that MAXV be weighted above IAMB (i.e., the requirement that feet be right headed) to guarantee that both syllables of /Sw/ words are preserved.

(8) Data from the Disyllabic Stage (Adam 2002: 79)

input: /…Sw/			input: /…wS/		
target	*child*	*gloss*	*target*	*child*	*gloss*
ˈse.fer	ˈse.fer	‘book’	ka.ˈdur	a.ˈdur	‘ball’
ʃar.ˈʃe.ret	ˈe.tet	‘necklace’	xa.ˈtul	a.ˈtul	‘cat’
la.ˈre.det	ˈde.det	‘to get down’	liʃ.ˈtot	i.ˈtot	‘to drink’
la.ˈka.xat	ˈka.xat	‘to take’	ne.ʃɪˈ.ka	ʃi.ˈka	‘kiss (n.)’

The only difference in weighting conditions necessary to distinguish the Trochaic Stage from the Disyllabic Stage is found in the relative weights of the Markedness constraint TROCHEE and the Faithfulness constraint MAXV. With the weight of TROCHEE reduced to below that of MAXV and the weight of ALIGN-FT-L remaining high, the disyllabic truncation pattern emerges. Pretonic syllables in target /…wSw/ forms continue to be deleted, leaving a single trochaic disyllable (9a). Pretonic syllables in /wS/ targets are now realized faithfully however (9b), resulting in the type of iambic disyllable not allowed during the earlier Trochaic Stage. Longer words with final stress are also truncated to disyllables – i.e., /wwS/ → [(wS)].

(9) a. Longer words with penultimate stress are truncated: /wSw/ → [(Sw)][4]

/ʃar.ˈʃe.ret/ ‘necklace’	ALIGN-FT-L *w*=100	MAXV *w*=60	TROCHEE *w*=40	*H*
☞ (ˈe.tet)		-1		-1(60) = **-60**
ar.(ˈe.tet)	-1			-1(100) = **-100**

(9) b. Disyllables with final stress are realized **faithfully**: /wS/ → [(wS)]

/ka.ˈdur/ ‘ball’	ALIGN-FT-L *w*=100	MAXV *w*=60	TROCHEE *w*=40	*H*
(ˈdur)		-1		-1(60) = **-60**
a.(ˈdur)	-1			-1(100) = **-100**
☞ (a.ˈdur)			-1	-1(40) = **-40**

The necessary weighting conditions for the Disyllabic Stage are contrasted with those for the Trochaic Stage in (10).

(10) Trochaic Stage: *w*ALIGN-FT-L, ***w*TROCHEE > *w*MAXV**
Disyllabic Stage: *w*ALIGN-FT-L > ***w*MAXV > *w*TROCHEE**

4 The choice to build a trochee rather than an iamb around the stressed syllable – i.e., (ˈe.tet) vs. *(a.ˈre) – reflects a well-known pressure in child grammars to retain final syllables (see e.g., Echols & Newport 1992, Pater 1997, Adam 2002). This pressure can be applied using an ANCHOR constraint (see esp. Pater 1997).

2.3 The Hebrew Intermediate-Faith Stage

Crosslinguistically, many children begin to acquire nouns before verbs[5] – a pattern robustly attested in Hebrew development (Adam & Bat-El 2007). In this context, Bat-El (2007) discusses a child whose lag in verb production is accompanied by an innovative pattern of morphologically-sensitive phonology. To quote Bat-El (2007: 8), "When the child starts producing verbs, he does not assume the prosodic structures already available for nouns, but rather starts fresh and proceeds on an independent (but not different) developmental path." While nouns are produced in accordance with the later Disyllabic Stage discussed in section 2.2, the newly-acquired verbs are produced in accordance with the earlier Trochaic Stage discussed in section 2.1. The contrast between these two patterns is apparent in target words with final stress.

(11) Disyllabic Stage (nouns): /...wS/ → [(**S**)]
Trochaic Stage (verbs): /...wS/ → [(**wS**)]

This difference in truncation patterns is based purely upon the morphological category of the target word. The grammar of the child includes both the Trochaic and Disyllabic truncation patterns, but their distribution is morphologically conditioned.

(12) Data from the Morphological Split (IF) Stage (Bat-El 2007)

	Nouns			**Verbs**		
Age	*target /...wS/*	***child [(wS)]***	*gloss*	*target /...wS/*	***child [(S)]***	*gloss*
1;05.04	bak.ˈbuk	buk.ˈbuk	'bottle'	maʃ.ˈpric	ˈpic	'squirt'
1;05.08	ar.ˈye	a.ˈye	'lion'	ho.ˈrid	ˈyit	'put down'
1;06.12	suf.gan.ˈja	in.ˈja	'donut'	ko.ˈfec	ˈfec	'jump'
1;06.20	liv.ya.ˈtan	i.ˈtan	'crocodile'	na.ˈfal	ˈfal	'fell'

Longitudinal data from Bat-El's (2007) study shows that this split in truncation behaviour persists, with some variation, for several months. The table in (13) tracks this development through the proportion of nouns and verbs with final stress that are realized as [(wS)] – i.e., in a manner consistent with the later Disyllable Stage – at each age.

(13) Longitudinal development of /…wS/ truncation patterns (Bat-El 2007)

	Nouns			**Verbs**		
Age Range	/…wS/ targets	[(wS)] outputs	% [(wS)] outputs	/…wS/ targets	[(wS)] outputs	% [(wS)] outputs
1;02.00-1;03.05	10	3	**30%**	–	–	–
1;03.14-1;05.14	45	32	**71%**	6	0	**0%**
1;05.08-1;07.02	244	187	**77%**	44	22	**50%**
1;07.09-1;09.00	346	296	**86%**	114	81	**71%**

[5] The scope of this generalization is a matter of some debate. See Gentner & Boroditsky (2001) for a recent review.

The developing phonology of this child thus shows three distinct stages – an early trochaic stage where almost exclusively nouns are produced and these are all reduced to trochees, a later intermediate stage where nouns are realized as trochaic or iambic disyllables and newly-acquired verbs are mostly realized as trochees, and a final stage where nouns and verbs are both generally realized as disyllables. Abstracting away from the variation, we can summarize the stages as in (14):

(14)

	a. Trochaic Stage	b. Intermediate Stage	c. Disyllabic Stage
	/…Sw/ → [(Sw)]	/…Sw/ → [(Sw)]	/…Sw/ → [(Sw)]
	/…wS/ → [(S)]	$/\ldots wS/_{noun}$ → [(wS)]	/…wS/ → [(wS)]
		$/\ldots wS/_{verb}$ → [(S)]	

In order to account for the difference in the truncation patterns of nouns and verbs at the innovative intermediate stage, a constraint is required that will protect segments in the privileged noun context from deletion, preferring the $/\ldots wS/_{noun}$ → [(wS)] mapping to one where the pre-tonic syllable is deleted. Drawing on work by Smith (2000, 2001), we propose that the relevant constraint is MAXV-NOUN, a constraint belonging to the family of positional faithfulness constraints (Beckman 1998).[6]

(15) MAXV-NOUN: *For nouns*, assign one violation mark to any vowel in the input that lacks an output correspondent.

As outlined above, the difference between the basic Trochaic and Disyllabic Stages comes from the relative weights of MAXV and TROCHEE (see 10). The addition of MAXV-NOUN allows us to capture the different treatment of finally-stressed nouns and verbs in this child's IF Stage through a *gang effect* of the type outlined in section 1.

(16) Weighting Conditions for Intermediate Faith Stage:
[*w*MAXV-NOUN + *w*MAXV] > *w*TROCHEE ⇒ Disyllabic pattern for nouns
*w*TROCHEE > *w*MAXV ⇒ Trochaic pattern for verbs

With ALIGN-FT-L heavily weighted and feet therefore necessarily left aligned, the weighting conditions in (16) produce the attested noun-verb asymmetry. As shown in (17), the inequality [*w*MAXV-NOUN + *w*MAXV] > *w*TROCHEE ensures that nouns with final stress retain two syllables in their output foot (17a), while the inequality *w*TROCHEE > *w*MAXV ensures that non-nouns with final stress are realized as monosyllabic trochees (17b).

(17) a. Finally-stressed **nouns** preserve two syllables: $/wS/_{noun}$ → [(wS)]

/ka.ˈdur/ 'ball'	TROCHEE *w*=60	MAXV *w*=40	MAXV-NOUN *w*=30	*H*
(ˈdur)		-1	-1	-1(40) + -1(30) = **-70**
☞ (ka.ˈdur)	-1			-1(60) = **-60**

---

6 See Bat-El (2007) for an alternative proposal based on preserving surface contrasts between morphological categories.

(17) b. Finally-stressed **verbs** are truncated to one syllable: $/\text{wS}/_{\text{verb}} \rightarrow [(\text{S})]$

/na.ˈfal/ 'fell'	TROCHEE *w*=60	MAXV *w*=40	MAXV-NOUN *w*=30	*H*
☞ (ˈfal)		-1		-1(40) = **-40**
(na.ˈfal)	-1			-1(60) = **-60**

This morphologically-sensitive stage is *intermediate* between the simple Trochaic and Disyllabic Stages discussed in sections 2.1 and 2.2. As the child passes from the Trochaic Stage to the IF Stage, the weights of general and specific faithfulness constraints increase until they are jointly great enough to gang up and overcome the weight of the conflicting markedness constraint. Then, as the child then passes from the IF Stage to the Disyllabic Stage, the weight of the general faithfulness constraint alone becomes great enough to overcome the weight of the conflicting markedness constraint. This trajectory of constraint inequalities is summarized in (18) below:

(18)	**Stage 1:**	*w*TROCHEE > **[***w*MAXV-NOUN + *w*MAXV**]** *w*TROCHEE > *w*MAXV	**Trochaic Stage**
	Stage 2:	**[***w*MAXV-NOUN + *w*MAXV**]** > *w*TROCHEE wTROCHEE > *w*MAXV	**IF Stage**
	Stage 3:	**[***w*MAXV-NOUN + *w*MAXV**]** > *w*TROCHEE *w*MAXV > *w*TROCHEE	**Disyllabic Stage**

In this section we have seen that IF stages like the Hebrew MAXV-NOUN case can emerge as children acquire their L1 phonology even when the target language provides no direct evidence for such a pattern. The following section demonstrates how the HG-GLA learner naturally passes through the sequence in (18) based solely on positive evidence from the target. This is a significant result, given the difficulties that ranked-constraint learners face when confronted with the same task (see, e.g., Hayes & Londe 2006, Tessier 2007a).

3. Gradual Learning and IF Stages

The HG-GLA learning procedure employed here shares two key properties with the OT Gradual Learning Algorithm (OT-GLA; Boersma 1998, Boersma & Hayes 2001). First, learning is *gradual*, proceeding in stage-like fashion from the initial state until a stable approximation of the target language is reached. Second, learning is *error-driven*, with the existing grammar being changed just when a target-language form cannot be faithfully reproduced. To these properties, we add the assumption that Markedness constraints are initially biased toward higher values than IO-Faithfulness constraints (cf. Demuth 1995, Gnanadesikan 2004, Pater 1997, Smolensky 1996). We implement this bias by starting the HG-GLA learning procedure with all Markedness constraints weighted at 100 and all IO-Faithfulness constraints weighted at zero (see Jesney & Tessier 2007 for discussion).

3.1 Gradual Learning with Weighted Constraints

With the constraints introduced in section 2 and a *w*M > *w*F initial state, our Hebrew learner naturally begins at the initial Trochaic Stage discussed in section 2.1. Both nouns and verbs with final stress are reduced to monosyllabic trochees, contrary to the pattern found in the target language. (In this section's tableaux, productions consistent with the Trochaic Stage pattern – i.e., errors – are indicated with 💣 and target-like disyllabic productions are indicated with ☞.)

(19) a. Stage 1: Target disyllabic **nouns** with final stress are truncated: /wS/ → [(S)]

/ka.ˈdur/ 'ball'	TROCHEE w=100	MAXV w=0	MAXV-NOUN w=0	*H*
💣 (ˈdur)		←● W_{-1}	←● W_{-1}	-1(0)+-1(0) = **0**
(ka.ˈdur)	L_{-1} ●→			-1(100) = **-100**

(19) b. Stage 1: Target disyllabic **verbs** with final stress are also truncated: /wS/ → [(S)]

/na.ˈfal/ 'fell'	TROCHEE w=100	MAXV w=0	MAXV-NOUN w=0	*H*
💣 (ˈfal)		←● W_{-1}		-1(0) = **0**
(na.ˈfal)	L_{-1} ●→			-1(100) = **-100**

From this initial stage, learning proceeds as in (20) (see Boersma & Weenink 2007, Jäger to appear, Pater to appear). This procedure ensures that, with each error, the values of constraints favouring the target output (the winner, W) are increased (indicated by ←●) and the values of constraints favouring the erroneous form (the loser, L) are decreased (indicated by ●→).

(20) Given a target input-output pair from the ambient language, the learner takes the input and feeds it to his grammar to determine which output candidate his current grammar deems optimal.

- If the output candidate selected by his grammar matches the target output, the grammar remains unchanged.
- If the output candidate selected by his grammar *does not* match the target output, each constraint value is the adjusted by $n(x - y)$, where $0 < n < 1$, and x is the number of violations of the constraint incurred by the loser and y is the number of violations of the constraint incurred by the winner.[7]

The errors in (19a) and (19b) both demonstrate that the loser-favouring constraint TROCHEE must be demoted. Errors involving nouns (19a) also demonstrate that the

[7] As Pater (to appear) notes, this readjustment procedure for weighted constraints is essentially the Perceptron update rule of Rosenblatt (1958). The general learning procedure is much the same as with the ranked-constraint GLA (Boersma 1998, Boersma & Hayes 2001); with the OT-GLA, however, all winner-favouring constraints are adjusted upward and all loser-favouring constraints are adjusted downward by the same small amount n, regardless of the number of violations incurred by the winner vs. the loser.

winner-favouring constraints MAXV and MAXV-NOUN must be promoted, while errors involving verbs (19b) provide additional evidence for the promotion of MAXV. Learning from these errors leads, after approximately 40 repetitions in Praat (Boersma & Weenink 2007), to the Intermediate Faithfulness Stage in (21) where finally-stressed nouns are realized faithfully as [(wS)] and finally-stressed verbs are truncated to [(S)].[8] With the adjusted weights, MAXV and MAXV-NOUN together can gang up to overcome the weight of TROCHEE, resulting in a disyllabic form for nouns (21a). The weight of MAXV alone, however, is insufficient to overcome the weight of TROCHEE, resulting in a trochaic monosyllabic form for verbs (21b).

(21) a. Stage 2: Target disyllabic **nouns** with final stress are not truncated: /wS/ → [(wS)]

/ka.ˈdur/ 'ball'	TROCHEE *w*=57	MAXV *w*=43	MAXV-NOUN *w*=21	*H*
(ˈdur)		-1	-1	-1(43)+ -1(21) = **-64**
☞ (ka.ˈdur)	-1			-1(57) = **-57**

(21) b. Stage 2: Target disyllabic **verbs** with final stress are truncated: /wS/ → [(S)]

/na.ˈfal/ 'fell'	TROCHEE *w*=57	MAXV *w*=43	MAXV-NOUN *w*=21	*H*
💣 (ˈfal)		←● W_{-1}		-1(43) = **-43**
(na.ˈfal)	L_{-1} ●→			-1(57) = **-57**

Since the IF grammar still makes errors with verbs (21b), the learner that has reached this stage continues to make adjustments to values of the implicated constraints TROCHEE and MAXV. Learning thus proceeds for approximately thirty more repetitions until reaching the Disyllabic Stage – which, for this grammar fragment, is the target. The HG-GLA learner has thus passed through each of the stages summarized in (18), reflecting precisely the patterns attested in child acquisition of Hebrew.

(22) Stage 3: Target disyllabic **verbs** with final stress are not truncated: /wS/ → [(wS)]

/na.ˈfal/ 'fell'	MAXV *w*=53	TROCHEE *w*=47	MAXV-NOUN *w*=21	*H*
(ˈfal)	-1			-1(53) = **-53**
☞ (na.ˈfal)		-1		-1(47) = **-47**

The key result of this section is that the HG-GLA learner, relying only on positive evidence from the target language, readily passes through the IF stage. Furthermore, this stage emerges without reference to cues in the learner's errors or to any *a priori* bias

---

[8] Praat simulation parameters: 'LinearOT' decision mode, 1.0 initial plasticity, 0.1 plasticity decrement, 2.0 evaluation noise. Nouns and verbs were fed to the learner in equal amounts; different proportions would produce different weights, but the gang effect IF stage will always emerge if the learner is exposed to at least some nouns.

affecting the weight of MAXV relative to that of MAXV-NOUN. As section 3.2 shows, this result is not as easily achieved with a grammar of ranked constraints.

3.2 Gradual Learning with Ranked Constraints

In order to model the three stages described above using a grammar of ranked OT constraints, the rankings in (23) must follow in sequence. Here, the IF Stage 2 is a consequence of the Specific Faithfulness constraint MAXV-NOUN outranking the Markedness constraint TROCHEE, which in turn outranks the General Faithfulness constraint MAXV.

(23)	**Stage 1:**	TROCHEE >> MAXV, MAXV-NOUN	**Trochaic Stage**
	Stage 2:	MAXV-NOUN >> TROCHEE >> MAXV	**IF Stage**
	Stage 3:	MAXV, (MAXV-NOUN) >> TROCHEE	**Disyllabic Stage**

As has been noted elsewhere (e.g., Hayes & Londe 2006, Tessier 2007a), the crucial IF Stage 2 does not follow simply in a gradual learning model like the GLA using ranked constraints (Boersma 1998, Boersma & Hayes 2001). The problem is tied to the superset-subset relationship that holds between the General and Specific Faithfulness constraints MAXV and MAXV-NOUN. The initial stage persists until one or both Faithfulness constraints outrank TROCHEE. Since errors involving both nouns and verbs prompt the promotion of MAXV but only errors involving nouns prompt the promotion of MAXV-NOUN, the first constraint to overcome TROCHEE will necessarily be the general MAXV, and the next stage will be that in (24) below.

(24) a. Stage 3: Disyllabic **nouns** with final stress are not truncated: /wS/ → [(wS)]

/ka.ˈdur/ 'ball'	MAXV	TROCHEE	MAXV-NOUN
(ˈdur)	*!		*
☞ (ka.ˈdur)		*	

(24) b. Stage 3: Disyllabic **verbs** with final stress are not truncated: /wS/ → [(wS)]

/na.ˈfal / 'fell'	MAXV	TROCHEE	MAXV-NOUN
(ˈfal)	*!		
☞ (na.ˈfal)		*	

The stage in (24) is the final Disyllabic Stage, not the attested Intermediate Faithfulness Stage that the gradual HG-GLA learner so easily passed through. Indeed, without additional assumptions that prove difficult to implement, the OT-GLA learner will *never* pass through a restrictive IF stage (for related discussion, see Hayes 2004, Prince & Tesar 2004, Tessier 2007a). While the difficulties facing the OT learner might eventually be overcome, we consider the HG-GLA learning results discussed here to be of particular importance precisely because these difficulties are never encountered. No additional biases or language-specific stipulations are required for the HG-GLA learner to pass through the full set of attested stages.

4. Conclusions

Children spontaneously innovate patterns of faithfulness in privileged contexts during the course of acquisition. Such stages are both *restrictive* – permitting marked structures only in a subset of the target environments – and *emergent* – observed in child language despite a lack of specific target-language evidence to support them.

Here we have considered a case of enhanced faithfulness to nouns over verbs in the acquisition of Hebrew. While this asymmetry has been previously reported in other adult phonologies (Smith 2000, 2001), we are not aware of any previous evidence for this IF stage in children. Previous work has, however, reported many IF stages involving prosodically-defined positions of prominence. Stressed syllables, for example, often admit more marked structures than unstressed syllables in developing grammars (in this context, see Tessier 2007a, and also Goad & Rose 2004, Revithiadou & Tzakosta 2004, Rose 2000). All such IF stages can be readily modeled using the HG-GLA learning procedure described here, but pose a considerable challenge for learning systems using ranked constraints (see Jesney & Tessier 2007, Pater, Jesney & Tessier 2007).

One interesting remaining issue is the provenance of the Specific Faithfulness constraints at work in IF stages. One possibility is that all constraints, including Specific Faithfulness constraints, are simply part of an innate constraint set. Another possibility is that the salience of particular elements in the input (nouns, onsets, stressed syllables, etc.) leads the learner to innovate faithfulness constraints specific to these categories and positions. Regardless of the source of these constraints, their co-existence with General Faithfulness constraints poses a serious challenge in modeling the acquisition of ranked constraint grammars. As this paper has shown, however, the co-existence of Specific and General Faithfulness constraints is unproblematic for a gradual learner using weighted HG constraints. Restrictive Intermediate Faithfulness stages emerge naturally in the HG-GLA model based simply upon the positive evidence provided by the target language.

References

Adam, Galit. 2002. *From Variable to Optimal Grammar: Evidence from Language Acquisition and Language Change*. PhD dissertation. Tel-Aviv University.

Adam, Galit & Outi Bat-El. 2007. Morphological knowledge without morphological structure: morphology-prosody interface in the first Hebrew verbs. To appear in *Generative Approaches to Hebrew Linguistics*, ed. S. Arnom-Lotem, G. Danon & S. Rothstein. [http://www.tau.ac.il/~obatel/adam_bat-el_morphological_knowledge.pdf].

Bat-El, Outi. 2007. Category-specific phonology in adult and child Hebrew. Handout from OCP4, Rhodes. [http://www.tau.ac.il/~obatel/bat-el_ocp4.pdf].

Beckman, Jill N. 1998. *Positional Faithfulness*. PhD dissertation. University of Massachusetts Amherst.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interactions Between Articulatory and Perceptual Drives.* PhD dissertation. University of Amsterdam.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.

Boersma, Paul & David Weenink. 2007. *PRAAT: Doing Phonetics by Computer*. [www.praat.org].

Demuth, Katherine & E. Jane Fee. 1995. Minimal words in early phonological development. Ms., Brown University and Dalhousie University.

Echols, Catharine H. & Elissa L. Newport. 1992. The role of stress and position in determining first words. *Language Acquisition* 2: 189-220.

Fikkert, Paula. 1994. *On the Acquisition of Prosodic Structure*. PhD dissertation. Leiden University.

Gentner, Dedre & Lera Boroditsky. 2001. Individuation, relativity and early word learning. In *Language Acquisition and Conceptual Development*, ed. M. Bowerman & S. Levinson, 215-256. Cambridge: Cambridge University Press

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

Goad, Heather & 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*, ed. René Kager, Joe Pater & Wim Zonneveld. Cambridge: Cambridge University Press.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the early stages. In *Fixing Priorities: Constraints in Phonological Acquisition*, ed. René Kager, Joe Pater & Wim Zonneveld. Cambridge: Cambridge University Press. [ROA-327].

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

Jäger, Gerhard. to appear. Maximum entropy models and stochastic Optimality Theory. In *Architectures, Rules, and Preferences: A Festschrift for Joan Bresnan*, ed. Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson & Annie Zaenen. Stanford: CSLI Publications.

Jesney, Karen & Anne-Michelle Tessier. 2007. Re-evaluating learning biases in Harmonic Grammar. In *UMOP 36: Papers in Theoretical and Computational Phonology*, ed. Michael Becker, 69-110. Amherst, MA: GLSA.

Keller, Frank. 2006. Linear Optimality Theory as a model of gradience in grammar. In *Gradience in Grammar: Generative Perspectives*, ed. Gisbert Fanselow, Caroline Féry, Ralph Vogel & Matthias Schlesewsky. Oxford: Oxford University Press.

Legendre, Géraldine, Yoshiro Miyata & 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, MA: Lawrence Erlbaum.

McCarthy, John J. & Alan Prince. 1993. Generalized alignment. *Yearbook of Morphology*, 79-153. [ROA-7].

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

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

Pater, Joe. to appear. Gradual learning and convergence. *Linguistic Inquiry*.

Pater, Joe, Karen Jesney & Anne-Michelle Tessier. 2007. Phonological acquisition as weighted constraint interaction. In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA 2)*, ed. Alyona Belikova, Luisa Meroni & Mari Umeda, 339-350. Somerville, MA: Cascadilla.

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

Prince, Alan. 2002. Anything goes. In *New Century of Phonology and Phonological Theory*, ed. Takeru Honma, Masao Okazaki, Toshiyuki Tabata & Shinichi Tanaka, 66-90. Tokyo: Kaitakusha. [ROA-536].

Prince, Alan S. & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell.

Prince, Alan S. & Bruce Tesar. 2004. Learning phonotactic distributions. In *Fixing Priorities: Constraints in Phonological Acquisition*, ed. René Kager, Joe Pater & Wim Zonneveld. Cambridge: Cambridge University Press. [ROA-353].

Revithiadou, A. & 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) 2003*. Utrecht: LOT Occasional Series.

Rose, Yvan. 2000. *Headedness and Prosodic Licensing in the L1 Acquisition of Phonology*. PhD dissertation. McGill University.

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*, ed. Rebecca Daly & Anastasia Riehl, 203-214. Ithaca, NY: CLC Publications.

Smith, Jennifer L. 2001. Lexical category and phonological contrast. In *PETL 6: Proceedings of the Workshop on the Lexicon in Phonetics and Phonology*, ed. Robert Kirchner, Joe Pater & Wolf Wikely, 61-72. Edmonton: University of Alberta. [ROA-728].

Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27: 720-731.

Smolensky, Paul & Géraldine Legendre. 2006. *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.

Tessier, Anne-Michelle. 2007a. *Biases and Stages in Phonological Acquisition.* PhD dissertation. University of Massachusetts Amherst. [ROA-883].

Tessier, Anne-Michelle. 2007b. Slowly learning to be faithful: frequency of violation and constraint-based learning. Ms., University of Alberta. [http://www.ualberta.ca/~annemich/Tessier2007b_UofAWPsDraft.pdf].

Karen Jesney
Department of Linguistics
226 South College
University of Massachusetts
Amherst, MA 01003 USA

kjesney@linguist.umass.edu

Anne-Michelle Tessier
Department of Linguistics
4-22 Assiniboia Hall
University of Alberta
Edmonton, AB T6G 2E7 Canada

amtessier@ualberta.ca