Testing for OO-Faithfulness in Artificial Phonological Acquisition

Anne-Michelle Tessier
University of Massachusetts Amherst

Abstract

This paper provides experimental evidence for the claims by Hayes (2004) and McCarthy (1998) that phonological acquisition relies on a bias for high-ranking paradigm uniformity (i.e. Output-Output Faithfulness). I report an experiment which tested the predictions of this proposed bias, using an artificial language 'wug test' (Berko, 1958) to test 4-year-old childrens' production of marked consonant clusters in two different morphological environments. The results support the claim that children prefer to repair clusters in ways that satisfy OO-Faith at the expense of other Markedness and Faithfulness pressures.

1. Theoretical Background: Ranking Biases and OO-Faith

Much of the OT literature in phonological acquisition, in both the theoretical learnability and experimental domains, has addressed the role of learning biases. In the terms of OT learnability, ranking biases are a theoretical mechanism to ensure that the learner adopts the most *restrictive* or conservative grammar consistent with the ambient data (i.e. to enforce some version of the Subset Principle -- see e.g. Berwick (1985), Wexler and Manzini (1987), Dresher and Kaye (1990).

The basic bias discussed in the OT learnability literature is Markedness >> Faithfulness (Smolensky, 1996; Ito and Mester, 1999; Tesar and Smolensky, 2000; Hayes, 2004; Prince and Tesar, 2004).[1] Empirically, the M >> F bias is supported by a body of experimental and corpus evidence suggesting that children begin with an unmarked grammar and learn to produce increasingly more marked structures over time, in domains including segmental inventory, phonotactics, syllable structure, and foot and prosodic structure (e.g. Demuth, 1995; Gnanadesikan, 1999/2004; Goad and Rose, 2004; Levelt and van de

* Thanks to Joe Pater, John McCarthy and Lyn Frazier for invaluable help; to Cushman School, Amherst, MA, for the support of its teachers, director, parents and enthusiastic children; and to Bruce Hayes, Kie Zuraw, Elan Dresher, John Kingston, Beth Hume, Michael Wagner, Kathryn Flack, some anonymous reviewers, and audiences at MOT 2005, HUMDRUM 2005, BUCLD30 and the University of Toronto for insights, challenges and discussion.

1. See Hale and Reiss (1998) for a dissenting view.

Vijver, 2004; Pater, 1997; Pater and Werle, 2001.) A second ranking bias already suggested in the literature, though not always in the present terms, is for Specific >> General faithfulness (Smith, 2000; Hayes, 2004; see also Reviathiadou & Tzakosta, 2004). Like M >> F, the bias for more specific faithfulness constraints is necessary to prevent the accidental acquisition of superset grammars. While Prince and Tesar (2004) raise serious issues for implementing the Specific-F >> General-F bias, I follow those authors who assume that positional faithfulness constraints are necessary to capture cross-linguistic generalizations (see esp. Beckman 1998) – and see Hayes (2004), Tessier (2006) for proposals of how this bias can be enforced.

This paper is an investigation of a third learning bias proposed in Hayes (2004) and McCarthy (1998): the preference for high-ranking Paradigm Uniformity constraints, formalized here using Output-Output Faithfulness (henceforth OO-Faith; Benua, 2000.). OO-Faith constraints require that derived words retain the phonological properties of their morphological bases.[2]

Two learnability arguments have been made for the high-ranking OO-Faith bias. One comes from McCarthy (1998), who points out that an inherent ranking of OO-Faith >> IO-Faith is necessary to learn OO-faithful languages that do not have surface alternations. In such a language, morphological bases are unmarked in some static way (e.g. they never violate word minimality) and this property holds of all members of the paradigm even when the markedness pressure on bases is not relevant to the derived forms (i.e., when derived forms are already big enough to avoid violating minimality.) McCarthy (1998) demonstrates using a Kansai dialect of Japanese that such a pattern can be understood as the effect of OO-Faith, and that acquiring the right restrictions on both simple and derived forms will require a bias for OO-faithfulness, since the lack of alternations will not provide any overt evidence for OO-faith's ranking.

The second learnability argument comes from Hayes (2004), who raises the issue of how children could acquire allophones whose phonological distribution is disturbed to keep a morphological paradigm uniform. The famous example he uses is the interaction of flapping and Canadian raising (CR) in some dialects of English. In such dialects, vowel raising is purely allophonic in mono-morphemic words: raised [ʌɪ] appears before voiceless obstruents as in 'write' [ɹʌɪt], and unraised [aɪ] appears elsewhere as in 'ride' [ɹaɪd]. However, derived forms with a base vowel [ʌɪ] exceptionally retain their raised quality even before a voiced flap, as in 'writer' [ɹʌiɾəɹ], *[ɹaiɾəɹ]. Hayes offers the OO-Faith bias as part of a learning model in which a child could learn [ʌɪ]'s distribution. When the learner realizes that 'writer' is derived from the base 'write' with raised [ʌɪ], OO-Faith constraints now prefer that same raised vowel in [ɹʌiɾəɹ]. A ranking bias to explain this vowel's distribution with OO-Faith rather than IO-Faith will thus guide the learner to the target ranking (see Hayes, 2004 for more of the details.)

---

2. Nothing crucial hinges on my choice of OO-Faith constraints to enforce paradigm uniformity; other proposals would change the details of the analyses but not the force of the arguments.

2. The OO-Faith Bias and Learning Predictions

Drawing together the arguments from the literature discussed above, we can express the full set of ranking biases in the unified constraint hierarchy of (1):

1) A full set of ranking biases
OO-Faith >> Markedness >> Specific-IO-Faith >> General-IO-faith

As assumed in the error-driven algorithms of Biased Constraint Demotion (Prince and Tesar, 2004), and Low-Faithfulness Constraint Demotion (Hayes, 2004), I take the biases in (1) to represent both the initial state of acquisition, as well as the learner's method of choosing between multiple possible rankings at every subsequent learning stage. (See the cited authors for the details of these models, which due to space constraints will simply be assumed here.)

2.1 The initial state

To see some of the predictions of (1), I consider the acquisition of one marked structure: obstruent clusters that disagree in voicing. Given (1), the initial grammar is one that prohibits any such clusters, because the markedness constraint Agree[voice] (Lombardi, 1999) ranks above all IO-Faithfulness constraints to voicing. The grammar in (2) shows this, using both a general Ident[voice], and a specific Ident constraint that targets onsets only (see e.g. Beckman 1998, Lombardi, 1999):

2) *The initial state: Agree[voice] is ranked high*[3]
Id[vce]-OO >> **Agree[voice]** >> $\text{Id[vce]}_{\text{Onset}}$-IO >> Id[vce]-IO

If the target language freely allows such clusters, however, the final state will be one in which Agree[voice] has been demoted below all Faith, i.e.:

3) *The final state: Agree[voice] has been demoted*
Id[vce]-OO >> $\text{Id[vce]}_{\text{Onset}}$-IO >> Id[vce]-IO >> **Agree[voice]**

To see the effects of OO-Faith, which necessarily requires morphological structure, we will consider each stages' optimal outputs for two lexical items with different morphology:

4) (a) /zɪtʃdɪn/ simple word
(b) /wʌtʃ + dəl/ derived word: morphological base /wʌtʃ/

---

3. In what follows here, the ranking between OO-Faith and Markedness is not crucial so long as they are both ranked above all IO-Faith. My choice of the ranking OO-Faith >> M is motivated by anecdotal data not discussed here – most notably Kazakis (1969) – but this choice requires more careful investigation.

Recalling that this is an idealization, and therefore assuming that the learner knows the morphological structure of their words: how will the initial stage grammar treat the two words in (4)? In simple forms like (4a), the grammar will repair the voicing mismatch of /tʃd/ by voicing the cluster's coda:

5) *Simple words at the initial state*

/zɪtʃdɪn/	Id-[vce]-OO	Agree[vce]	Id[vce]$_{Onset}$-IO	Id-[vce]-IO
zɪtʃdɪn	N/A	*!		
zɪtʃtɪn			*!	*
☞ zɪdʒdɪn				*

In derived forms like (4b), the repair is different: coda voicing is protected by OO-Faith as part of a morphological base, so voicing is repaired in *onset*:

6) *Complex words at the same initial state*

/wʌtʃ + dəl/	Id[vce]-OO	Agree[vce]	Id[vce]$_{Onset}$-IO	Id[vce]-IO
wʌtʃdəl		*!		
☞ wʌtʃtəl			*	*
wʌdʒdəl	*!			*

2.2 Beyond the initial state

Since learners do not immediately leap to the correct final state grammar: how does the grammar change as learners move past the initial state? Without the space here to argue for an explicit model of how intermediate stages can be generated (the subject of Tessier, in prep.), it seems in keeping with the spirit of learning via constraint demotion to assume an intermediate stage below[4]:

7) Id[vce]-OO >> Id[vce]$_{Onset}$-IO >> **Agree[voice]** >> Id[vce]-IO

Compared to the initial state in (2), the ranking (7) has conceded to the data by demoting Agree[voice] below Ident-Onset, but keeps to its biases by retaining Agree[voice] above general Ident. At this stage, what outputs are now optimal for our two-word lexicon? The repair in simple forms has not been affected:

8) *Simple words at an intermediate stage:*

/zɪtʃdɪn/	Id[vce]-OO	Id[vce]$_{Onset}$-IO	Agree[vce]	Id[vce]-IO
zɪtʃdɪn	N/A		*!	
zɪtʃtɪn		*!		*
☞zɪdʒdɪn				*

4. One well-known OT learning model that does produce intermediate stages is the Gradual Learning Algorithm (Boersma and Hayes, 2001.) One crucial question for future research is how reliably the GLA learns when faced with constraints like ID-ONSET.

In the derived form, however, the effect of this intermediate re-ranking is now to block *all* possible repairs for cluster voicing mismatches. For the form /wʌtʃ + dəl/, OO-Faith blocks voicing change in coda, and onset-specific IO-Faith blocks a repair in onset. As a result, the optimal candidate is the target one, in which mid-ranking Markedness is violated and the voicing mismatch survives:

9) *Complex words at an intermediate stage:*

/wʌtʃ + dəl/	Id[vce]-OO	Id[vce]$_{Onset}$-IO	Agree[vce]	Id[vce]-IO
☞ wʌtʃdəl			*	
wʌtʃtəl		*!		*
wʌdʒdəl	*!			*

This mini-example has yielded two different stages predicted by the OO-Faith- biased learner, which both show asymmetries between the treatment of simple and derived words. The experiment described in the rest of this paper investigated whether such stages can be induced in an experimental context.

3. Experiment: Testing for the influence of OO-Faith

The initial and intermediate stages of the previous section are intended to be predictions about the nature of children's L1 learning. However, investigating such questions in natural L1 learning poses some large potential confounds. For one thing, the predicted effects of OO-faith at any stage are intrinsically tied to the learner's representational assumptions about morphology – its bases, relations, paradigms, and the like – and childrens' acquisition and interpretation of morphology may vary considerably. Furthermore, English morphology does not provide many good testing grounds for such investigation.

Thus, the experiment I report below was an artificial language learning study, in which children learned both novel coda-onset clusters as well as a novel plural suffix in an 'alien' language invented by the experimenter. This design was an attempt to simulate a morphologically-informed initial state – and one in which we can be sure that participants were encountering the language's bases and suffix for the first time.

3.1 Experimental predictions

The words of the artificial language taught in this experiment were designed to compare faithfulness to marked coda-onset clusters in two morphological conditions: count nouns like [wʌtʃ.del], where the first syllable was the singular base and the second syllable was the novel plural suffix, and mass nouns like [zɪtʃ.dɪn], where both syllables were neither base nor affix.

10) *Prediction 1, with respect to initial syllable codas*

Coda segments should be produced more faithfully in the first syllable of plural nouns than in the first syllable of mass nouns – e.g., more faithfulness to /tʃ/ in /wʌtʃ/+ /dəl/ than in /zɪtʃdɪn/

Once children have learned the relevant morphology, this asymmetry should hold under both the initial and intermediate rankings. In both cases, base codas should be protected – either at the expense of onset segments at the initial state, or Markedness violations at the intermediate state – whereas mass noun first-syllable codas should not.

11) *Prediction 2, with respect to unfaithful medial clusters:*
Among those tokens whose medial clusters are produced unfaithfully in some way, more of the unfaithfulness should be seen in onset position in the count nouns than the mass nouns.

This prediction is a specific test for the asymmetry of the initial state. Where high-ranking Markedness has driven an unfaithful repair, the ranking of OO >> IO faith predicts *onset* repairs for count nouns, but not for mass ones.

3.2 Methodology and Materials

Twelve 4-year old children in the Amherst and Northampton, Massachusetts areas participated in the study. The experiment was presented to the children as a novel language-learning game, in which their task was to learn some "alien" words, spoken by an alien puppet named Bozdim (operated by the experimenter.) Children were taught the alien words by association with pictures of familiar objects: count nouns in singular and plural contexts, and mass nouns in two different containers. The children therefore learned a series of new nouns, as well as the novel plural suffix [dəl].

In training, the experimenter (and puppet) presented children with picture one at a time. First the child was asked to tell Bozdim the English name for the object; then they were asked the puppet what the object was called in his language. The child was encouraged to repeat the object's name ("Can you say what Bozdim just said?"), and to use it in discussion of the object ("Is this a blue wutch? Or is it a yellow wutch?" or "Does it look like this cup of zitchdin tastes good? Do you think the zitchdin is hot or cold?") Each child first learned three words of one noun class – count or mass – and then three of the other (the . Within a count noun block, participants first learned three singular nouns, and then their corresponding three plurals. Within a mass noun block, participants learned three mass nouns, and then heard the same three again in a different container (e.g., a glass of juice, and later a bottle of juice.) Half of the participants saw count nouns before mass nouns, and the other half saw mass before count.

Count noun singulars were all mono-syllables, of the shape CVC(C); each singular was suffixed with [dəl] to form a bisyllabic plural with the shape CVC(C).[dəl]. Mass nouns were all bisyllabic, of the shape CVC(C).dVC. All bisyllabic forms were initially stressed; all vowels in the second syllable were lax (of the set ɪ, ɛ, ə) and pronounced as unstressed but not completely reduced. Every effort was made by the experimenter to produce the clusters and their segments similarly in all tokens and contexts: in particular, coda stops were somewhat released (as they might be in very careful English speech). A sample initial training block is given in (12):

12) *A sample training block* [5]

Noun class	*Morphology*	*Prosodic shape*	*Target words*	*Matching pictures*
Count	Three singulars (base)	CVC(C)	[pob]	one armchair
			[wʌtʃ]	one pick-up truck
			[nænf]	one flower
	Three plurals (base + suffix)	CVC(C).dəl	[pobdəl]	many armchairs
			[wʌtʃdəl]	many pick-up trucks
			[nænfdəl]	a garden of flowers
Mass	Three mass nouns	CVC(C).dVC	[gɪbdɛt]	a glass of juice
			[zɪtʃdɪn]	a cup of cocoa
			[gʌnfdɛp]	a mug of milk
	Same three mass nouns	CVC(C).dVC	[gɪbdɛt]	a bottle of juice
			[zɪtʃdɪn]	lots of cups of cocoa
			[gʌnfdɛp]	a carton of milk

Once children had learned three words from both categories, the experimenter asked the children to play a matching game with the puppet. All twelve pictures seen so far were laid out in front of the child, and the puppet pointed to one picture and named it. The child would then find the matching picture, and name it for the puppet. In the game, the puppet pointed to one of each of the mass nouns, for the child to match by naming the other, and to each of the *singular* count nouns, for the child to match by naming the *plural*. Thus, each child was asked to provide six words, all with difficult coda-onset clusters: three mass nouns, and three plural nouns derived from singular bases.

5. In addition to these clusters, 5 others were used but that did not yield useful data. They were excluded from analysis either because they were not pronounced unfaithfully in more than 2 tokens (clusters: g.d, k.d) or because they were not produced by more than 1 child in both the mass and plural contexts (clusters: ft.d, kt.d, bʒ.d).

After the first game, children were presented another training block like the one in 12), with four more words, and another testing game was played.

4. Results

To make claims about the role of OO-Faith in the participants' productions, we must be able to claim that participants had in fact learned the artificial language's morphology – that is, learned its plural suffix "del". In order to prove sufficient mastery of /dəl/, I required that participants provide at least one spontaneous token of more than one plural noun, associated with the right plural picture. This criterion eliminated 2 participants, leaving 10 children. All results reported are for the 5 clusters and 10 children, across all tokens produced in the course of the experiment. One exception is that plural tokens were only included when the participant produced a second syllable of type dV(C): tokens with English plural affixes ("wutʃez, pobdelz") or zero morphology ("wutʃ, pob") are not included.

4.1 General results

The majority of children's pronunciations were of two types: either faithful, or with reduction in the coda of the first syllable. To first give an impression of the data, table 13 summarizes the general results (variances in parentheses):

13) *Results, across subject and by condition*

	total tokens	*faithful codas out of total tokens*		*unfaithful medial clusters out of total*		*faithful codas out of total unfaithful clusters*	
	#	#	%	#	%	#	%
plural	87	70/87	0.793 (0.035)	25/87	0.287	9/25	0.36 (0.1711)
mass	112	56/112	0.50 (0.06)	52/112	0.464	1/52	0.019 (0.006)
totals	199	126		77		10	

4.2 Testing prediction 1

The data in table 14 below allows us to test prediction 1 – that codas in initial syllables should be more faithful in count nouns than mass ones. The table shows the raw number of faithful coda productions by subject and conditions, and also the proportion of all tokens that were coda-faithful in each condition:

14) *Proportion of faithful σ1 codas (Cs) by subject and condition*

	Codas of Mass Nouns			*Codas of Plural Nouns*		
Subj.	*# faithful*	*# total*	*%C-faith*	*# faithful*	*# total*	*%C-faith*
C	17	20	**0.85**	11	11	**1**
E	4	9	**0.444**	9	9	**1**
A2	2	7	**0.286**	6	6	**1**
I	4	15	**0.267**	5	8	**0.625**
N2	8	12	**0.667**	9	12	**0.75**
A3	1	10	**0.1**	7	9	**0.778**
A1	3	9	**0.333**	10	15	**0.667**
D1	13	17	**0.765**	3	4	**0.75**
D2	1	3	**0.333**	7	10	**0.7**
N1	3	10	**0.30**	3	3	**1**
totals	**56**	**112**		**70**	**87**	
means			**0.5**			**0.805**
variance			0.06			0.024

Summing across all 10 subjects, a two-tailed t-test showed that codas were produced faithfully significantly less often in mass noun clusters, namely in 50% of tokens, than in the plural count noun clusters, where they were faithful 80.4% of the time ($p < 0.01$). Further, a pair-wise two-tailed t-test, comparing the proportion means for each subject, also shows a significant difference between the lower proportions of faithful first-syllable codas produced in mass nouns compared to the higher proportions in plural nouns ($p < 0.01$).[6] Thus, prediction 1 seems nicely borne out. This result provides some evidence for intermediate stage rankings, where plural nouns are faithful to both members of medial clusters, but mass nouns are still unfaithful in coda position.

4.3 Testing prediction 2

Table 15 below tests prediction 2 – that derived forms will show not only more faithfulness in their cluster codas, but also less faithfulness in cluster onsets. To test for such a possibility, I consider again the proportion of faithful codas, but only among those that were unfaithful somewhere in the medial cluster:

6. Statistics were calculated both for all 10 subjects, but also using just the first 8, since the low number of total items for the last two subjects, D2 and N1, might have skewed the proportions. Either way, however, the result is significant at $p < 0.01$.

15) *Proportion of faithful codas in unfaithful medial clusters*

	Codas Mass Nouns			*Codas in Plural Nouns*		
Subj.	*# faithful*	*# total*	*%C-faith*	*# faithful*	*# total*	*%C-faith*
C	1	4	**0.25**	0	0	**0**
E	0	5	**0**	1	1	**1**
A2	0	3	**0**	0	0	**0**
I	0	8	**0**	0	3	**0**
N2	0	4	**0**	1	4	**0.25**
A3	0	9	**0**	3	5	**0.6**
A1	0	6	**0**	1	6	**0.1667**
D1	0	4	**0**	0	1	**0**
D2	0	2	**0**	0	3	**0**
N1	0	7	**0**	0	**0**	**0**
totals	**1**	**52**		**6**	**23**	
means			**0.0192**			**0.261**
variance			0.0063			0.1711

Table 15) shows that unfaithful clusters were overwhelmingly unfaithful in coda position: only 1/52 mass nouns and 9/25 plural nouns with unfaithful clusters had faithful codas. This is not too surprising, given the privileged faithful status of onsets over codas. Despite this clear tendency to apply repairs in coda, table 15) does provide some support for prediction 2. Across all 10 subjects, a one-tailed t-test assuming unequal variances shows that the mean proportion of coda faithfulness in these clusters is lower for mass than for plural nouns ($t = -2.077$; $p = 0.0322$.[7]

5. General Discussion

The experiment reported here provides novel experimental evidence that children rely on high-ranked OO-Faith constraints in phonological acquisition. When 4-year-old children were faced with novel difficult consonant clusters in a new language, their repairs to those clusters demonstrated preferences for OO-faithfulness over both Markedness and IO faithfulness.

In the present experiment, one large and unanswered question is why children are *ever* unfaithful to initial syllable codas in plurals? According to the theory outlined here: once children have learned the suffix ‘del’: both the initial and intermediate state rankings protect base material at the expense of something else (either faithfulness in affixes or else markedness) and therefore a

7. Across just the first 8 subjects, the difference between the means remains significant with a one-tailed t-test ($t = -2.2131$; $p = 0.0289$).

plural noun's initial syllable coda should remain untouched. This prediction is clearly too strong for my results: table (16) in particular showed that while base codas were more faithful than other codas, they were still much *less faithful than onsets* overall. One answer may lie with the mental resources required to implement an OO-faithful grammar: setting up a lexical entry for a closed class affix like 'plural', constructing a morphologically-complex input to the phonological grammar and the like. It is remains unknown, at least in this methodology, how and when the morphological knowledge that [dəl] is an affix was used online, either to prompt learning via constraint re-ranking or to simply rule out suboptimal candidates that violated OO-Faith. There are many such questions to be addressed: all revolving around the question of what kind of phonological knowledge or grammar is being used or constructing by children in this experimental paradigm. Another set of theoretical and methodological questions that remain for future work concern the connection between these artificial learning results and the nature of real-life L1 acquisition.

Taking a step back from these issues, however, this experiment's positive result is encouraging in at least two ways. First, it provides support for any learning theory which predicts the early emergence of paradigm uniformity, even in the absence of such evidence in the ambient language. Second, it provides an indication that children are both willing and able to engage in an experiment of this type, particularly in learning new functional material like a plural suffix. This suggests that artificial language learning may be a valuable methodology for tapping children's grammatical development, and especially their sensitivity to patterns and distinctions not found in their L1.

References

Beckman, Jill N. (1998) *Positional faithfulness,* Ph.D. dissertation, UMass Amherst. Amherst, MA: GLSA Publications.

Benua, Laura (2000). *Phonological Relations between Words.* Outstanding Dissertations in Linguistics, Garland Publishing.

Berko, Jean G. (1958) "The child's learning of English morphology," *Word* 14: 150-177.

Berwick, Robert (1985) *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

Boersma, P. and B. Hayes (2001) "Empirical tests of the gradual learning algorithm," *Linguistic Inquiry* 32: 45-86.

Demuth, Katherine (1995) "Markedness and the development of prosodic structure," in J. Beckman (ed.) *Proceedings of NELS 25*, Amherst MA: GLSA Publications.

Dresher, B. Elan. and Jonathan Kaye (1990) "A Computational Learning Model for Metrical Phonology," *Cognition* 34.2: 137-195.

Gnanadesikan, Amahlia (2004) "Markedness and faithfulness constraints in child phonology," in R. Kager, W. Zonneveld, J. Pater (eds.), *Fixing Priorities: Constraints in Phonological Acquisition,* Cambridge: Cambridge University Press.

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 R. Kager, W. Zonneveld, J. Pater (eds.), *Fixing Priorities: Constraints in Phonological Acquisition,* Cambridge: Cambridge University Press.
Hale, Mark and Charles Reiss (1998) "Formal and empirical arguments concerning phonological acquisition," *Linguistic Inquiry* 29: 656-683.
Hayes, Bruce (2004) "Phonological Acquisition in Optimality Theory: the early stages," in R. Kager, W. Zonneveld, J. Pater (eds.), *Fixing Priorities: Constraints in Phonological Acquisition,* Cambridge: Cambridge University Press.
Ito, Junko and Armin Mester (1999) "The phonological lexicon," in N. Tsujimura (ed.), *The Handbook of Japanese Linguistics.* Oxford: Blackwell.
Kazazis, Kosta (1969) "Possible evidence for (near)-underlying forms in the speech of a child," in R. Binnick et al, (eds.), *Proceedings of the Chicago Linguistic Society 5,* Chicago: CLS.
Levelt, Claartje C. and Ruben van de Vijver (2004) "Syllable types in cross-linguistic and developmental grammars," in R. Kager, W. Zonneveld, J. Pater (eds.), *Fixing Priorities: Constraints in Phonological Acquisition,* Cambridge: Cambridge University Press.
Lombardi, Linda (1999) "Positional faithfulness and voicing assimilation in Optimality Theory," *Natural Language and Linguistic Theory* 17.2: 267 – 302.
McCarthy, John J. (1998) "Morpheme structure constraints and paradigm occultation," in M. Catherine Gruber et al (eds.), *Proceedings of the Chicago Linguistic Society 32, Vol. II: The Panels,* Chicago, CLS.
Pater, Joe (1997) "Minimal Violation and Phonological Development," *Language Acquisition* 6*:* 201-253.
Pater, Joe and Adam Werle (2001) "Typology and variation in child consonant harmony," in C. Féry, A. D. Green and R. van de Vijver, (eds.), *Proceedings of HILP 5*, University of Potsdam.
Prince, Alan and Paul Smolensky (2004) *Optimality Theory: Constraint interaction in generative grammar*, Oxford: Blackwell.
Prince, Alan and Bruce Tesar (2004) "Learning Phonotactic Distributions," in R. Kager, W. Zonneveld, J. Pater, (eds.), *Fixing Priorities: Constraints in Phonological Acquisition,* Cambridge: Cambridge University Press.
Revithiadou, Anthi and Marina Tzakosta (2004) "Markedness Hierarchies vs. Positional Faithfulness and the Role of Multiple Grammars in the Acquisition of Greek," *Proceedings of Generative Approaches to Language Acquisition (GALA)* 2003, Utrecht: LOT Occasional Series.
Smolensky, Paul (1996) "On the comprehension/production dilemma in child language," *Linguistic Inquiry* 27: 720-731.
Smith, Jennifer L. (2000) "Positional faithfulness and learnability in Optimality Theory," in Rebecca Daly and A. Rehl, (eds.), *Proceedings of ESCOL99,* Ithaca: CLC Publications.
Tesar, Bruce and Paul Smolensky (2000) *Learnability in Optimality Theory*, Cambridge, MA: MIT Press.
Tessier, Anne-Michelle (2006) Learning Stringency Relations and the Contexts of Faithfulness. Talk presented at the 80th LSA meeting, Albuquerque, New Mexico.
Tessier, Anne-Michelle (in prep.) *Biases and Stages in OT Phonological Acquisition.* Ph.D. dissertation, UMass Amherst. Amherst, MA: GLSA Publications.
Wexler, Kenneth and M. Rita Manzini (1987) "Parameters and learnability in binding theory," in T. Roeper and E. Williams, (eds.), *Paramater Setting,* Dordrecht: Reidel**.**