

Positional Faithfulness and Learnability in Optimality Theory

Jennifer L. Smith, University of Massachusetts

1. Introduction: Toward a learnable theory of phonology and typology

In Optimality Theory (Prince & Smolensky 1993), the grammar of a language consists of a set of constraints and a ranking of those constraints. The constraint set is universal,¹ so the only language-particular aspect of a grammar is its constraint ranking. As a result, constraint ranking plays a crucial role in many domains of language. For example, the **phonology** of a language (its phonotactic patterns and allophonic alternations) is the result of one particular constraint ranking. **Typological patterns** across languages reflect all possible constraint rankings. **Learning** a language requires learning its constraint ranking.

Because phonology, typology, and learning all depend on constraint ranking, each of these three domains can contribute to our understanding of the others. Any new constraint that is proposed for the phonological analysis of one language must also make correct typological predictions — every possible reranking of that constraint must correspond to some possible human language. Furthermore, if the phonological analysis proposed for a given language includes a ranking that is not formally learnable, then that analysis must be discarded, since no language that corresponds to an unlearnable ranking can exist.

There are many examples of OT analyses in which typological facts are used to refine the formulation of the constraints developed to account for some phonological phenomenon. One such proposal is the theory of positional faithfulness (Beckman 1995, 1998; Casali 1996), summarized in section 2 below. Positional faithfulness constraints are formulated in a particular way because certain (logically possible) phonological patterns are not observed in any language, so there must be no ranking of the constraints that is able to generate those patterns.

This paper examines positional faithfulness from the perspective of learnability (as formulated in the Error-Driven Constraint Demotion learning algorithm of Tesar & Smolensky 1996, 1998, Smolensky 1996). Since positional faithfulness is a theory that takes into account both phonological phenomena and typological patterns, the question addressed here is the following: What can be learned about the grammar when phonology, typology, and learnability are considered simultaneously? It is shown below that a grammar with a positional faithfulness ranking is formally learnable only if a condition is placed on the ranking of constraints in the learner's Initial State. Namely, the set of positional faithfulness constraints must initially be ranked above the set of general faithfulness constraints.²

2. Positional faithfulness

The theory of positional faithfulness (Beckman 1995, 1998; Casali 1996) is one

approach that has been taken within OT toward the problem of positional neutralization (Trubetsky 1939), a type of phonological process in which material that is contained in a 'strong' position is resistant to neutralization processes (generally featural) that affect material in the corresponding 'weak' position.

(1) Examples of positional neutralization patterns

<i>Position with contrast</i>	<i>Position with neutralization</i>	
Roots	Affixes	(McCarthy & Prince 1995)
Initial syllables	Non-initial σ	(Selkirk 1994; Beckman 1995; Casali 1996)
Stressed syllables	Unstressed σ	(Selkirk 1994; Alderete 1995)
Released consonants ³	Unreleased C	(Kingston 1985; Lombardi 1991; Steriade 1993, 1997)

A crucial part of accounting for positional neutralization is finding a way to capture the typological asymmetry of the neutralization. For a given featural contrast (e.g., having mid vowels as distinct phonemes), (a) some languages have the contrast in all positions (e.g., Japanese); (b) some languages have the contrast in no positions (e.g., Quechua); and (c) some languages have positional neutralization, with the contrast in a strong position only (e.g., Russian, in stressed syllables). However, no language has the contrast in a weak position only. Therefore, if constraints are proposed in order to account for positional neutralization phenomena, they must be designed so that no constraint ranking corresponds to a language that maintains a featural contrast in a weak position only.

One solution that has been given for the problem of positional neutralization is the following theory of positional faithfulness (Beckman 1995, 1998).⁴ In this theory, each faithfulness constraint in the grammar has a specific version relativized to every strong position. (The inclusion of a particular position in the class of 'strong positions' is justified when that position has intrinsic psycholinguistic or phonetic salience; see Beckman 1998 for discussion.) For example, there is a positional version of the faithfulness constraint IDENT[Vht] 'Output forms maintain input specifications for vowel height' for the strong position *stressed syllable* ($\acute{\sigma}$). A language like Russian, that maintains mid vowels only in stressed syllables, would have the following constraint ranking.

(2) Ranking for a language with mid vowels in stressed syllables only

IDENT/ $\acute{\sigma}$ [Vht] >> *MIDV >> IDENT[Vht]
└─ Contrast in $\acute{\sigma}$ ─┘
└─ No contrast ─┘ (for the language in general)

As outlined in Prince & Smolensky (1993), contrast is determined by the relative ranking of markedness constraints (such as *MIDV) and faithfulness constraints (such as IDENT constraints; McCarthy & Prince 1995). **M** >> **F** means that contrast is prohibited, while **F** >> **M** means that contrast is preserved. Therefore, the ranking

in (2) ensures that contrastive mid vowels are banned except in stressed syllables.

This positional faithfulness approach is able to capture the asymmetrical nature of positional neutralization, as seen in (3), where all permutations of the ranking of markedness, faithfulness, and positional faithfulness (**PosF**) constraints are given.

(3)	<i>Ranking</i>	<i>Characteristic</i>	<i>Result</i>
a.	M >> PosF >> F	(M highest ranked)	No contrast in any position
b.	M >> F >> PosF		
c.	PosF >> F >> M	(F >> M)	Contrast possible in all positions
d.	F >> PosF >> M		
e.	F >> M >> PosF		
f.	PosF >> M >> F		Contrast in strong position only

As this ranking permutation demonstrates, no ranking of these constraints can produce a language with a contrast in weak positions only. This result is brought about by Beckman's (1998) crucial restriction on positional faithfulness constraints: they are only available for strong positions. If a faithfulness constraint specific to a weak position, such as IDENT/ǫ[Vht], were included in the universal constraint set, then the theory would be forced to predict the existence of languages that rank that constraint as in (3f) above. Such a ranking would allow contrastive mid vowels in unstressed syllables only — a pattern that never occurs. Therefore, constraints like IDENT/ǫ[Vht] are explicitly disallowed. It is in this sense that positional faithfulness is a phonological theory that has been shaped by typological considerations.

As noted in the introduction, it is not only phonological alternations and typological patterns that interact by means of constraint ranking. In grammar learning, given that the set of constraints is universal, it is precisely the constraint ranking that must be learned. This paper explores the three-way relationship among phonology, typology, and learnability by asking the following questions. First, if the theory of positional faithfulness as outlined above is taken to be part of the grammar, then is this kind of grammar compatible with considerations of formal learnability? And if so, what can be learned about positional faithfulness or about grammar learning from examining the interaction between the two? The following section introduces the learning algorithm that will be applied to positional faithfulness in section 4 in order to address these questions.

3. Grammar learning as Error-Driven Constraint Demotion

The learning algorithm known as Error-Driven Constraint Demotion (EDCD; Tesar & Smolensky 1996, 1998; Smolensky 1996; see also Tesar 1999 and Prince & Tesar 1999 for further developments) has been proposed as a theory of how a constraint ranking can successfully be learned in an OT grammar.⁵ This section gives a brief

overview of EDCD and its implications for the structure of the learner's Initial State ranking; the interaction of EDCD and positional faithfulness is discussed in the following section.

3.1. The EDCD algorithm

The basic idea behind EDCD is that the learner gradually approaches the target (adult) grammar by demoting constraints that are seen to be violated in adult output forms. For example, suppose that the learner is at an intermediate stage of acquisition, with a constraint ranking as in (4).

- (4) The learner's current, non-target ranking
 $\{ *MIDV, \dots, \dots \} \gg \{ IDENT[Vht], \dots, \dots \}$

The learner then observes a target-language form (an actual adult output) of the shape [tep]. Taking that adult output as its input,⁶ the learner computes the optimal output that its current constraint ranking would produce.

- (5) Computing the output of the current grammar

<i>Input:</i> /tep/	*MIDV	IDENT[Vht]
a. tep	*!	
☞ b. tip		*

If the learner's predicted output matches the actual form, no action takes place; EDCD is 'error-driven' because constraint demotion only occurs when, as in (5), the learner's predicted output fails to match the target form. When this kind of mismatch does occur, demotion proceeds as follows. First, the uncanceled constraint violations (known as "marks") of the two forms are compared. An uncanceled violation, or mark, is a violation by one form that is not matched by a violation of the same constraint by the other form.

- (6) Comparison of marks

	<i>Actual output:</i> [tep]	<i>Current prediction:</i> [tip]
<i>Marks:</i>	*MIDV	IDENT[Vht]

Constraints are then demoted so that all marks from the actual adult form are dominated by at least one mark from the currently predicted but incorrect output.

- (7) Constraint demotion

~~*MIDV~~ >> IDENT[Vht] >> *MIDV

Next, the learner proceeds to evaluate adult forms with the new grammar resulting from the constraint demotion. When the learner reaches a stage in which the correct adult form is chosen as the output of its own grammar every time, then the target language has been learned.

Tesar & Smolensky (1996, 1998) demonstrate that in order for the EDCD learning algorithm to converge on target grammars correctly, demotion must be carried out conservatively: the constraint being demoted remains as high in the hierarchy as possible while still allowing the observed adult form to be more harmonic than the output form produced by the pre-demotion grammar. In terms of the example given above, *MIDV is demoted to a rank just below IDENT[Vht], not to a rank arbitrarily low in the hierarchy. However, the process of demotion is not permitted to impose any rankings between two constraints that are not being demoted. Thus, in demoting constraint A below constraint B, if B is still unranked with respect to a constraint C, then A is demoted below the unranked constraint "stratum" consisting of { B C }.

EDCD is like many other theories of acquisition and learnability in that it does not require the learner to have access to negative evidence. Under EDCD, constraint demotion only occurs when there is *positive evidence*, namely, an observed form that the learner's current grammar does not generate.

3.2. *The Initial State ranking*

The EDCD algorithm is a theory of how the learning of a grammar proceeds from the learner's initial state until the target constraint ranking is reached. However, the algorithm itself does not specify what constraint ranking the learner has at the start of the learning process. The nature of this Initial State ranking is a question that must also be addressed.

The hypothesis requiring the least imposition of structure on the Initial State is that initially, all constraints are unranked. However, Smolensky (1996) has demonstrated that this hypothesis cannot be correct. If all constraints are initially unranked, then there is danger that the learner might encounter a version of the Subset Problem (Berwick 1985, Wexler & Manzini 1987): a situation in which the learner's current grammar is incorrect, but the forms produced by the target grammar are a proper subset of the forms predicted by the current, incorrect grammar. In such a case, all adult forms that the learner encounters are compatible with its current grammar. Since learning (here, constraint demotion) occurs only when there is positive evidence that the current grammar does not match the target grammar, no learning can take place, and the target grammar can never be reached.

To see this problem, consider the following situation. The target language has no mid vowels, which means that the target constraint hierarchy includes the ranking *MIDV >> IDENT[Vht]. Now suppose the learner has a current ranking that includes IDENT[Vht] >> *MIDV. The learner will never encounter an adult form with a mid vowel, because the target language does not allow mid vowels. However, all the adult forms that the learner encounters will be compatible with its current ranking.

(8) The Subset Problem

- a. Target grammar: *MIDV >> IDENT[Vht]
- b. Current grammar: IDENT[Vht] >> *MIDV
- c. Observed adult form: [tip] ([tup], [tap], etc.)
- d. Prediction by current grammar: [tip] ([tup], [tap], etc.)

<i>Input:</i> /tip/	IDENT[Vht]	*MIDV
a. tep	*!	*
☞ b. tip		

- e. Action: *none*

Because all observed forms are compatible with the current grammar, no demotion takes place; the learner's ranking must remain IDENT[Vht] >> *MIDV. However, this is the wrong ranking for a language that permits no mid vowels. Since OT allows no language-particular restrictions on input forms (according to the principle of "Richness of the Base"; Prince & Smolensky 1993), all phonotactic patterns in a language must result from its constraint hierarchy. Therefore, a language with no mid vowels must have as part of its grammar a ranking that forbids them, namely, *MIDV >> IDENT[Vht].⁷ But this means that the learner in (8) can never reach the correct grammar for the language that it is trying to learn.

Smolensky (1996) shows that the only way that the learner is guaranteed to reach a target ranking like *MIDV >> IDENT[Vht] is by starting from that ranking. If it is the correct ranking for the target language, then by EDCD, no change will take place, and the target language will have been appropriately learned. Furthermore, starting from this same Initial State ranking also allows the learning of a target language with the opposite ranking, IDENT[Vht] >> *MIDV. If the initial *MIDV >> IDENT[Vht] ranking is not correct for the target language, then the target language is one that has mid vowels. As seen in (4)-(7) above, positive evidence in the form of adult outputs with mid vowels will always allow the learner to demote *MIDV below IDENT[Vht]. Therefore, both possible rankings of these two constraints are learnable if the Initial State ranking is *MIDV >> IDENT[Vht].

In general, an adult grammar that lacks a marked feature (e.g., contrastive mid vowels) must have the ranking **M** >> **F**. A grammar with **M** >> **F** produces a subset of the forms of a grammar with **F** >> **M**. Therefore, if the learner starts from a ranking with **F** >> **M**, it can never reach a target grammar with **M** >> **F**; there will be no positive evidence that its current ranking is incorrect, so no reranking can take place. However, if the Initial State has **M** >> **F**, a target grammar with **F** >> **M** can always be learned, because adult forms violating **M** are available to provide positive evidence that **M** must be demoted. So an Initial State ranking of **M** >> **F** is the only ranking that ensures that target grammars of both types, **M** >> **F** and **F** >> **M**, can be learned. Smolensky (1996) thus proposes that the Initial State ranking is structured

as follows.

$$(9) \quad \{ \mathbf{M} \} \gg \{ \mathbf{F} \}$$

That is, all markedness constraints are unranked with respect to one another, and all faithfulness constraints are unranked with respect to one another, but all **M** dominate all **F**. This Initial State ranking is needed to avoid the Subset Problem.

The proposed Initial State ranking is also compatible with findings from child phonology (e.g., Gnanadesikan 1995, Demuth 1995). The phonological behavior of children during acquisition does indeed show evidence of an initial high rank for **M** constraints even when they are crucially dominated in the target adult language.

4. Positional faithfulness and learnability

In the preceding sections, this paper has summarized two theories, positional faithfulness and the EDCD learning algorithm:

Positional faithfulness is a theory designed to analyze a phonological phenomenon, namely positional neutralization, in which typological considerations have influenced the formulation of the constraints. This theory holds that the strong positions in the grammar have special positional faithfulness (**PosF**) constraints. If a language has the ranking **PosF** \gg **M** \gg **F**, then it has positional neutralization; the contrast banned by **M** is absent in the language in general, but present in the strong position *Pos*.

The EDCD theory of learning holds that the Initial State ranking is $\{ \mathbf{M} \} \gg \{ \mathbf{F} \}$. Learning proceeds by constraint demotion, but constraints are demoted only in response to positive evidence.

Now that these background concepts have been introduced, the core question can be addressed: What happens when a grammar that makes use of positional faithfulness (i.e., a theory that is concerned with phonological processes and typological patterns) is the target grammar for learning by the EDCD algorithm? More specifically, are these two theories compatible, and if so, what are the implications for the grammar when they are taken together?

The rest of this section provides a demonstration of how the EDCD algorithm fares when faced with a language that has positional neutralization and the corresponding **PosF** \gg **M** \gg **F** ranking. It is shown that unless the Initial State ranking has **PosF** constraints dominating **F** constraints, the very same kind of Subset Problem arises that was demonstrated for **M** and **F** by Smolensky (1996).

4.1. Learning a grammar with positional faithfulness (first attempt)

Consider the case of a target language that, like Russian, has mid vowels only in stressed syllables. The target ranking that the learner must acquire is that in (10).

$$(10) \quad \text{IDENT}/\acute{\sigma}[\text{Vht}] \gg * \text{MIDV} \gg \text{IDENT}[\text{Vht}]$$

If the Initial State ranking for grammar learning is $\{ \mathbf{M} \} \gg \{ \mathbf{F} \}$ as in Smolensky (1996), then the constraints relevant for the current discussion are ranked in the Initial State as follows.

(11) $\{ *MIDV, \dots, \dots \} \gg \{ IDENT[Vht], IDENT/\acute{O}[Vht], \dots \}$

Suppose the learner encounters the form [tépa]. With this form as its input, the current grammar produces the output [típa].

(12) Output of current grammar

<i>Input: /tépa/</i>	*MIDV	IDENT[Vht]	IDENT/acute{O}[Vht]
a. tépa	*!		
b. típa		*	*

This output differs from the observed adult output, so the marks incurred by the two forms are compared and constraint demotion takes place.

(13) Application of EDCD

a. Comparison of marks

Marks: *Actual output:* [tépa] *Current prediction:* [típa]
 *MIDV IDENT[Vht], IDENT/acute{O}[Vht]

b. Constraint demotion

$$\overbrace{*MIDV \gg \{ IDENT[Vht], IDENT/\acute{O}[Vht] \} \gg *MIDV} \leftarrow$$

(Although the comparison of marks indicates that *MIDV would only need to be demoted below one of the two faithfulness constraints, the faithfulness constraints have not been ranked with respect to one another, so *MIDV is demoted to a position immediately below the stratum containing the two unranked faithfulness constraints.)

At this point, in order to reach the target grammar in (10) above, the learner would have to continue by demoting the positional faithfulness constraint IDENT/acute{O}[Vht] below the markedness constraint *MIDV. But the discussion in section 3.2 has shown that the learner will never encounter positive evidence that causes a faithfulness constraint to be demoted below a markedness constraint. That is, exactly the same type of Subset Problem has arisen here.

4.2. Learning a grammar with positional faithfulness (second attempt)

Just as the solution to Smolensky's (1996) original Subset Problem lay in fixing the Initial State ranking of **M** constraints with respect to **F** constraints, the solution to this new Subset Problem is another specification of ranking relationships in the Initial

State. If all **PosF** constraints initially outrank all general **F** constraints, then a grammar with the positional faithfulness ranking **PosF** >> **M** >> **F** becomes formally learnable.

Assume the same target language as in section 4.1 above — a language with mid vowels in stressed syllables but not in other positions. The target ranking for this language is thus still that in (10), repeated below in (14).

$$(14) \quad \text{IDENT}/\acute{\sigma}[\text{Vht}] \gg *MIDV \gg \text{IDENT}[\text{Vht}]$$

However, now the Initial State ranking has { **M** } >> { **PosF** } >> { **F** }, giving rise to the following initial ranking for the relevant constraints.

$$(15) \quad \{ *MIDV, \dots, \dots \} \gg \{ \text{IDENT}/\acute{\sigma}[\text{Vht}], \dots, \dots \} \gg \{ \text{IDENT}[\text{Vht}], \dots, \dots \}$$

For the adult form [tépa], the output of the learner's current grammar is [típa], again leading to mark comparison and constraint demotion.

(16) Output of current grammar

<i>Input:</i> /tépa/	*MIDV	IDENT/acute sigma[Vht]	IDENT[Vht]
a. tépa	*!		
☞ b. típa		*	*

(17) Application of ED CD

a. Comparison of marks

Actual output: [tépa] *Current prediction:* [típa]
Marks: *MIDV IDENT/acute sigma[Vht], IDENT[Vht]

b. Constraint demotion

$$\begin{array}{c}
 \cancel{*MIDV} \gg \text{IDENT}/\acute{\sigma}[\text{Vht}] \gg *MIDV \gg \text{IDENT}[\text{Vht}] \\
 \underbrace{\hspace{10em}} \\
 \phantom{\cancel{*MIDV}} \gg \text{IDENT}/\acute{\sigma}[\text{Vht}] \gg *MIDV \gg \text{IDENT}[\text{Vht}]
 \end{array}$$

This time, the learner reaches the target grammar. No Subset Problem arises.

The difference between this attempt and the previous unsuccessful one is that now, the Initial State ranking contains an additional ranking specification among the faithfulness constraints such that { **PosF** } >> { **F** }. As a result, when the learner observes an adult form that violates **M** in a strong position, it is not forced to demote **M** below both **PosF** and **F**. Instead, demoting conservatively, it demotes **M** only below **PosF**. The further demotion of **M** below general **F**, needed for a language that allows contrast in all positions (not only in strong positions), is of course possible, but it requires positive evidence in the form of an **M** violation outside a strong position.

Both of these Initial State ranking restrictions — the { **PosF** } >> { **F** }

ranking proposed here, and the $\{ \mathbf{M} \} \gg \{ \mathbf{F} \}$ ranking of Smolensky (1996) — allow the learner to avoid the Subset Problem because they provide a bias toward unmarked, subset grammars by encouraging \mathbf{M} constraints to remain ranked above as many $(\mathbf{Pos})\mathbf{F}$ constraints as possible. \mathbf{M} will end up below \mathbf{F} only when absolutely necessary, under compulsion from positive evidence. This strategy is further developed by Prince & Tesar (1999), who propose a mechanism by which \mathbf{M} constraints can remain as high as possible, and \mathbf{F} constraints as low as possible, not only in the Initial State, but throughout the course of learning.

5. Conclusions

Two results have emerged from this discussion. First, it has been demonstrated that grammars with positional faithfulness constraints are learnable by the Error-Driven Constraint Demotion algorithm. Therefore, these two theories are compatible as parts of a complete theory of grammar. (Of course, merely showing that the two theories are compatible does not prove that they are both correct. However, if they had been incompatible, then at least one of them would have been incorrect.)

Second, observing how the EDCD handles a target grammar with the positional neutralization ranking of $\mathbf{PosF} \gg \mathbf{M} \gg \mathbf{F}$ has shown that if positional faithfulness and the EDCD learning algorithm are both part of UG, there must be an additional specification on the Initial State ranking such that $\{ \mathbf{PosF} \} \gg \{ \mathbf{F} \}$.

In general, examining how phonology, typology, and learning interact can provide information about how the grammar is structured. This interconnectedness is one of the attributes of Optimality Theory that makes it a promising avenue of research. Because claims about one area of the grammar (i.e., typology) have implications for another (i.e., grammar learning), it is very often possible to find support (or disconfirmation) for a particular proposal in another domain of the grammar. As a result, many hypotheses are externally testable.

Endnotes

*Many thanks to John McCarthy, Alan Prince, Bruce Tesar, and audiences at the 1999 Rutgers-UMass Joint Optimality Workshop and ESCOL 99 for comments and discussion; any errors or inadequacies are my responsibility. This research was partially supported by the National Science Foundation under grant SBR-9420424 and by an NSF Graduate Research Fellowship.

¹A number of researchers have proposed that Alignment-type constraints may refer to specific morphemes (e.g., EDGEMOST in Prince & Smolensky 1993:35). If this is so, then part of the universal formulation of Alignment constraints is that they have the ability to take individual morphemes of a language as their "arguments."

²Hayes (1999) reaches similar conclusions; Prince & Tesar (1999) discuss these and other implications for the learning of constraints that stand in a 'specific-to-general' relation as positional faithfulness and general faithfulness constraints do. An analogous fixed initial ranking between output-output and input-output faithfulness

constraints is proposed by McCarthy (1999).

³ Although 'syllable onset' is commonly identified as a strong position, Steriade (1993, 1997) has convincingly shown that the consonants that resist neutralization of, e.g., laryngeal and place features are not those that are onsets, but those that are *released* (and therefore contain the best cues for the recovery of the contrast in question).

⁴ Other accounts of positional neutralization in OT include a different conception of positional faithfulness that requires a universally fixed ranking (Casali 1996, McCarthy & Prince 1995); positional markedness (Steriade 1997, Zoll 1998); and an alignment-based approach (Zoll 1997, 1998).

⁵ Other recent work on formal learnability in phonology includes Dresher (1999), Dresher & Kaye (1990), Hale & Reiss (1998), Hayes (1999), and Pulleyblank and Turkel (1997).

⁶ As noted by, e.g., Pulleyblank & Turkel (1997) and Hayes (1999), the earliest phonological acquisition by children appears to proceed without taking morphophonological relationships into account. Therefore, they argue, it seems reasonable to assume that, in early acquisition, input forms closely resemble output forms, and the learner has no access to paradigm-based alternations.

⁷ Empirical evidence supporting this claim, that a language lacking a certain marked structure in output forms corresponds to a constraint ranking that formally bans that structure, can be found in loanword phenomena. Generally, if a language bans mid vowels, then monolingual speakers of that language will (at least initially) actively alter the mid vowels that appear in words borrowed from other languages.

References

- Alderete, J. 1995. "Faithfulness to prosodic heads." Ms., University of Massachusetts.
- Beckman, J. 1995. "Shona height harmony: Markedness and positional identity." J. Beckman et al., eds., *UMOP 18*. Amherst: GLSA, 53-76.
- Beckman, J. 1998. *Positional Faithfulness*. PhD dissertation, University of Massachusetts.
- Berwick, R. 1985. *The Acquisition of Syntactic Knowledge*. Cambridge: MIT Press.
- Casali, R. 1996. *Resolving Hiatus*. PhD dissertation, UCLA.
- Demuth, K. 1995. "Markedness and the development of prosodic structure." J. Beckman, ed., *Proceedings of NELS 25*. Amherst: GLSA.
- Dresher, B.E. 1999. "Charting the learning path: Cues to parameter setting." *Linguistic Inquiry* 30.
- Dresher, B.E., and J. Kaye. 1990. "A computational learning model for metrical phonology." *Cognition* 34:137-195.
- Gnanadesikan, A. 1995. "Markedness and faithfulness constraints in child phonology." To appear in a volume from Cambridge University Press. Available on the Rutgers Optimality Archive [<http://rucss.rutgers.edu/roa.html>], #67.

- Hale, M., and C. Reiss. 1998. "Formal and empirical arguments concerning phonological acquisition." *Linguistic Inquiry* 29:656-683.
- Hayes, B. 1999. "Phonological acquisition in Optimality Theory: The early stages." To appear in a volume edited by R. Kager and W. Zonneveld. Available on the Rutgers Optimality Archive, #327.
- Kingston, J. 1985. *The Phonetics and Phonology of Timing between Oral and Glottal Events*. PhD dissertation, University of California, Berkeley.
- Lombardi, L. 1991. *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts. [New York: Garland Press, 1994.]
- McCarthy, J. 1999. "Morpheme structure constraints and paradigm occultation." M.C. Gruber et al., eds., *CLS 32, vol. II: The Panels*. Chicago: Chicago Linguistic Society.
- McCarthy, J., and A. Prince. 1995. "Faithfulness and reduplicative identity." J. Beckman et al., eds., *UMOP 18*. Amherst: GLSA, 249-384.
- Prince, A., and P. Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms., Rutgers University and University of Colorado, Boulder.
- Prince, A., and B. Tesar. 1999. "Learning phonotactic distributions." Ms., Rutgers University. Available on the Rutgers Optimality Archive, #353.
- Pulleyblank, D., and W.J. Turkel. 1997. "Gradient retreat." I. Roca, ed., *Derivations and Constraints in Phonology*. Oxford: Clarendon Press, 153-193.
- Selkirk, E. 1994. "Optimality Theory and featural phenomena." Lecture notes, Linguistics 730, University of Massachusetts.
- Smith, J. 1999. "Theories of faithfulness: Implications for learnability." Handout from presentation at the Fourth Rutgers-UMass Joint Optimality Workshop, Rutgers University, New Brunswick.
- Smolensky, P. 1996. "The Initial State and 'Richness of the Base' in Optimality Theory." Technical Report JHU-CogSci-96-3, Johns Hopkins University. Available on the Rutgers Optimality Archive, #154.
- Steriade, D. 1993. "Positional neutralization." Handout from presentation at NELS 24, University of Massachusetts, Amherst.
- Steriade, D. 1997. "Phonetics in phonology: The case of laryngeal neutralization." Ms., UCLA.
- Tesar, B., and P. Smolensky. 1996. "Learnability in Optimality Theory (long version)." Technical Report JHU-CogSci-96-4, Johns Hopkins University. Available on the Rutgers Optimality Archive, #156.
- Tesar, B., and P. Smolensky. 1998. "Learnability in Optimality Theory." *Linguistic Inquiry* 29, 229-268.
- Trubetsky, N. S. 1939 [1969]. *Principles of Phonology*. C.A.M. Baltaxe, trans. Berkeley/Los Angeles: University of California Press.
- Wexler, K. and Manzini, M.R. 1987. "Parameters and learnability in Binding Theory." T. Roeper and E. Williams, eds., *Parameter Setting*. Boston: Reidel.
- Zoll, C. 1997. "Conflicting directionality." *Phonology* 14:2, 263-286.
- Zoll, C. 1998. "Positional asymmetries and licensing." Ms., MIT.

jlsmith@linguist.umass.edu