How to Parse (and How not to) in OT Phonology

Mark Hale & Charles Reiss

Concordia University

1. How Not to Parse in OT

1.1 The claims

Smolensky (1996, henceforth S) attempts to account for the peculiarities of children's speech output and the well-known discrepancy between their inaccurate production of adult words and their extremely accurate parsing of adult speech by appealing to the state of their grammar. In particular S rejects the notion that there is a "dramatically greater performance/competence gap for children" (p.1). S proposes that a single OT grammar can generate both adult-like comprehension and child-like production if one assumes that at the initial state of the grammar, S_0, OT structural, or Well-formedness (W) constraints are ranked above Faithfulness (F) constraints which value correspondence between input forms (underlying representations) and output forms (surface phonetic representations). S's proposal is represented in (1):

(1) Single OT grammar generates adult-like comprehension and degenerate, unmarked child output if at S_0:
Wellformedness constraints >> Faithfulness constraints

1.2 The argument

The key to S's argument lies in drawing a distinction between the nature of production and comprehension in an OT model. This distinction is sketched in (2) and (3):

(2) PRODUCTION: OT-grammar selects the most 'harmonic' *output/surface form* (from the set of candidates which GEN provides) for a given *input/UR*

(3) COMPREHENSION: (same) OT-grammar's selects the most harmonic *input/UR* for a given observed *output/surface form*

K. Kusumoto (ed.), NELS 27, 159-170

As S states, "What differs between "production and 'comprehension' is only *which structures compete*: structures that share the same underlying form in the former case, structures that share the same surface form in the latter case" (p.3).

The result of making such a distinction is that the two operations will not always lead to the same input-output mapping, e.g. at S_0. In (4) we have adapted S's constraint tableaux to show how the distinction works. Compare the pronunciation of a stored lexeme /kæt/ to the comprehension of this same lexeme as pronounced by an adult. Following S we ignore irrelevant details of the pronunciation such as the aspiration on the initial voiceless stop. Since the W-constraints are all ranked high, every possible candidate form except for the most unmarked will violate some W-constraints. Like S, we have not distinguished among candidates on the basis of which constraints they violate, since this does not affect the structure of S's valid argument. Again following S, we assume that the universally least marked output representation is [ta]. Since this candidate violates no W-constraints, it is selected by the grammar at this stage as the optimal surface form. Note that the same candidate will surface no matter what input form is used at this stage of the grammar. In the bottom half of the tableau we illustrate how, in S's system, the child is able to parse adult [kæt] accurately as /kæt/ using the same grammar. Since the surface form [kæt] is a given, the mapping from any possible underlying representation to this surface form will violate the same W-constraints. The surface form is known *a priori* to violate constraints against the presence of a coda, of an [æ] and of a dorsal consonant. Therefore, it is left to the F-constraints to select the most harmonic, the optimal, input-output mapping.

(4) The grammar at the initial state following Smolensky 1996

- PRODUCTION: /kæt/ pronounced [ta] ('emergence of the unmarked')
- COMPREHENSION: [kæt] parsed as /kæt/, not [skæti], since mapping of /kæt/ to [kæt] is more harmonic than /skæti/ to [kæt] (only F-constraints matter).[1]

	Candidates	W-constraints (*æ, *Dorsal, *Coda...)	F-constraints (Parse, Fill...)
PRODUCTION			
UR /kæt/			
☞	[ta]		*
	[kæt]	*!	
	[skæti]	*!	*
	[dajpərræʃ]	*!	*
	etc.	*!	*
COMPREHENSION			
Surface [kæt]			
	/ta/	*	*!
➔	/kæt/	*	
	/skæti/	*	*!
	/dajpərræʃ/	*	*!
	etc.	*	*!

The winner in comprehension is marked with an arrow, ➔.

Before moving on to the problem with S's proposal, note that S assumes (correctly, we believe) that children store URs fully and accurately specified, according to what they hear in the target language.[2] Therefore, under S's own analysis the notion of *richness of the base* (e.g. Prince & Smolensky 1993, 191) becomes, irrelevant (at least in the case of non-alternating forms) to the acquisition process. Richness of the base is a claim about the nature of OT grammars which suggests that there can be great latitude in the form of URs. For example, the lexicon of English could have all voiceless velar stops stored as clicks, but given the appropriate constraint ranking, e.g. with constraints against clicks ranked high, the surface forms would still be pronounced with normal velar stops.

[1] S p.7: "What is given is the surface form, so the competing structures now [i.e. in comprehension—mrh&cr] are all those which are *pronounced* [kæt]." This is uninterpretable for child language as stated, since, by S's hypothesis, the grammar is responsible initially for maximally 'unmarked' pronunciation. There is no UR at this stage which is *pronounced* [kæt], even though adult [kæt] can be parsed as such. It seems clear from the discussion that S is trying to say that the mappings from every possible UR to surface candidate [kæt] are compared. The UR corresponding to the most harmonic mapping is the winner.

[2] We would qualify this by allowing for errors in parsing, which lead to incorrect representations. This is to be distinguished from merely incomplete representations assumed by researchers who posit that the child does not have access to all the features of the universal feature inventory. See Hale and Reiss 1996ad for arguments against this position. Since S denies the relevance of performance in the characterization of language acquisition, he cannot appeal to such parsing errors. This idealization does not, however, affect the structure of his argument, which up to this point we accept.

But given S's own assumptions about how parsing and the acquisition of URs proceeds, there will never be any reason to expect that /kæt/, for example, would be stored with a click. This computational curiosity of OT grammars, while real, thus may to be totally irrelevant to human language.

1.3 The problem

We now turn to a demonstration of the major flaw in S's argument. Since the parsing (comprehension) algorithm that S proposes (for both children and adults) generates the most harmonic mapping from a UR to a surface form, the algorithm will never be able to account for the well attested and widespread phenomenon of surface ambiguity, or merger, in natural language. German, for example, has two surface forms [rat], one derived from the UR /rat/ and the other from /rad/. We can capture the phenomenon of coda devoicing in German by assuming that a constraint against voiced codas is ranked above constraints demanding faithfulness to underlying voicing values. The relevant aspects of German grammar are sketched in (5).

(5) German (or any other) surface ambiguity
- /rat/ > *rat* 'advice' /rad/ > *rat* 'wheel'
- *VoicedCoda >> Faith[Voice]

Consider what happens when a surface form [rat] is parsed by a speaker of German, using S's algorithm. Since the surface form is a given in parsing, and since the choice of UR is left to the F-constraints, the most harmonic mapping from a UR to [rat] will be from the UR /rat/. The mapping from UR /rad/ to surface [rat] violates the same W-constraints as the mapping from /rat/ to [rat], but the former violates more F-constraints than the latter. This is, of course, a general result: in any case of surface merger, only the most 'unmarked' underlying lexeme will be chosen by the parse, since this lexeme provides the most faithful mapping.

2. Some Proposals on How to Parse in an OT Grammar

It is clear that S's parsing algorithm must be replaced with one which generates a *set* of parses, not a single parse, if we are to account for surface ambiguity. We propose two such algorithms for parsing a surface form Φ. In (6) we sketch an algorithm which is in the non-procedural spirit of OT. Under the assumption that massive computational complexity will ultimately be amenable to effective modeling, the algorithm culls the set of all possible URs to select those which can serve as a parse for a given surface form.

(6) 'Shrinking' algorithm in the 'spirit of OT':
To select a set of possible parses for a surface form Φ: (a) GEN generates all possible Urs: Ψ_i, i=1, . . . ; (b) for each UR Ψ_i GEN generates all possible surface candidates; (c) for each UR Ψ_i whose optimal output is Φ, Ψ_i is a parse for Φ.

In (7) we sketch a more procedural algorithm which starts with a set of parses containing only the one form which is identical to the surface form. The algorithm expands the hypothesis space of the parse by 'undoing' the effects of W-constraints.

(7) 'Expanding' algorithm

- Let the set of possible parses for Φ, $I = \{\Psi_i, i=1, \ldots\}$ be equal to Φ; $I = \{\Phi\}$
- Start at the highest ranked constraint;
- When a F-constraint which refers to a feature G is encountered 'fix' the candidate set with respect to G. That is, all subsequent candidates must be identical to some Ψ_i with respect to the feature G.
- When a W-constraint is encountered, expand candidate set along precisely the dimension specified by the W-constraint. I.e. add candidates Ψ_i to the hypothesis space which differ from some preexisting candidate only in violating the current W-constraint.
- The algorithm ends when there is no remaining W-constraint which dominates a F-constraint. The parse candidate set thus produced $I = \{\Psi_i, i=1, \ldots, k\}$ represents the set of URs which will be neutralized to Φ by the grammar.

We can illustrate the operation of the algorithm in (7) by contrasting the parsing of English [rat] vs. [rad] with that of the ambiguous German [rat], assuming the URs in (8).

(8) Contrastive parsing

• English	/rat/ 'rot'	/rad/ 'rod'
• High German	/rat/ 'advice'	/rad/ 'wheel'.

Since English does not have a rule of coda devoicing, we can assume that the ranking of *VoicedCoda in English is the opposite of that assumed for German, above. The operation of the parsing algorithm is sketched in (9) where a single UR is associated with surface [rat].

(9) English parse of [rat]: Faith[voice] >> *VoicedCoda

- The candidate set consists of /rat/
- The voicing specification of all segments in /rat/ is fixed by Faith[voice]
- The candidate set is not increased by *VoicedCoda , since [voice] has been 'fixed' in previous step
- The overt form is associated to a *single* UR, /rat/.

In German, on the other hand, the algorithm leads to an ambiguous parse, as desired, shown in (10).

(10) German parse of [rat]: *VoicedCoda >> Faith[voice]

- The candidate set consists of /rat/
- The candidate set is expanded to /rat/ and /rad/
- Voicing specification of all segments in /rat/ and /rad/ fixed by Faith[voice]
- The overt form is ambiguous—derivable from both /rat/ and /rad/

Whichever algorithm turns out to be better, it is obvious that either of our proposals is superior to S's, since they generate a *set* of candidate URs for a given surface form. Note that the argument developed here for phonology applies to S's syntactic example as well.

Sentences with differing underlying structures can be ambiguous on the surface. Clearly, a theory of sentence comprehension must provide a mechanism for generating such ambiguity.

3. Implications for the initial ranking of Faithfulness constraints

We now turn to a consideration of the implications of having a parsing algorithm that works (i.e. generates a *set* of underlying forms) for the study of the learnability of OT grammars. There is, first of all, an intuitive argument to be made against the position held by S and virtually every other scholar writing about the learning of OT grammars. Since surface forms and underlying forms tend to be 'fairly close' in adult grammars, it is clear that most F-constraints must ultimately be ranked higher than W-constraints. A theory which assumes that the F-constraints start out ranked high is preferable *a priori* to one which posits massive reranking.

This intuitive argument can, however, be supported by a demonstration that a parsing algorithm that actually works **requires** that F-constraints be initially ranked high in UG so that learners can converge on a lexicon. In contrast to S, then, we propose that the initial state of the grammar must be that shown in (11).

(11) At S_0: Faithfulness constraints >> Wellformedness constraints

With the initial ranking proposed in (11) there is a single outcome to each parse at S_0. With the initial ranking proposed by S in (1) a parsing algorithm like (6), which eliminates candidates from an initially infinite set, will generate the empty set; and one like (7), which adds candidates to an initially unary set, will explode the candidate set to include all possible URs. A lexicon is unacquirable under either scenario.

The table in (12) illustrates the acquisition of English /rat/ and /rad/ (forms AB) as opposed to German /rat/ and /rad/ (forms C-F), based on exposure to relevant surface forms. The German forms ending in [-əs] are genitive singular forms of the relevant nouns; because the stem-final stops occur between vowels, i.e. in onset position, in these forms, coda devoicing cannot apply. In the top half of the table we sketch the learning path under the assumption that all F-constraints are ranked high. Using either parsing algorithm, (6) or (7), the learner will be able to converge on a single UR for each surface form. Using (6), the high ranking of all F-constraints ensures that the optimal candidate is identical to the input form. Using (7), the high ranking of all F-constraints 'fixes' the value of all features of the surface form before the W-constraints can expand the set of candidate parses, again producing a single, fully faithful parse at the initial state.

The parse chosen is the correct one with respect to the adult grammar in each case except for form E. Ultimately, when the grammar generates the alternations due to coda devoicing, forms E and F will have to be collapsed. This process is obviously intimately related to the process of constraint reranking, whereby *VoicedCoda is raised above Faith[Voice] to obtain the grammar of German.

(12) Comparing HiFaith and LoFaith at S_0 using a parser that works

Surface form		Initial Hypothesis for UR	Path to adult UR
With F constraints ranked HIGH			
A.	rat	rat	Unique, correct UR is selected initially
B.	rad	rad	"
C.	rat	rat	"
D.	ratəs	rat	"
E.	rat	rat	E & F stored differently, later collapsed
F.	radəs	rad	by storing /rad/ and raising *VoicedCoda
With F constraints ranked LOW			
A.	rat	∅ or rat, rat⊙ʂ, bəbə...	There can be no
B.	rad	∅ or rat, rat⊙ʂ, bəbə...	learning path: each production yields the
C.	rat	∅ or rat, rat⊙ʂ, bəbə...	maximally unmarked utterance, say *ta*, as
D.	ratəs	∅ or rat, rat⊙ʂ, bəbə...	S desires, but each parse yields ∅ by (8)
E.	rat	∅ or rat, rat⊙ʂ, bəbə...	or else everything generated by the
F.	radəs	∅ or rat, rat⊙ʂ, bəbə...	UG-given W-constraints by (9).

The bottom half of the table illustrates the problem with assuming that F-constraints are initially ranked low. As desired by S, the production mapping will generate the maximally unmarked [ta] at the initial state, but (6) will generate no parses—there is no UR which will surface as [rat] at this stage, since every UR will surface as [ta]—and (7) will generate an infinite set of candidate parses, since no features of the surface form Φ will be 'fixed' before the W-constraints expand the parse set to include forms with every possible W-constraint violation.

It is worth reiterating at this point that the reranking of constraints and the collapsing of predictable allomorphs to a single form are two aspects of a single process, despite the following suggestions to the contrary:

(13) Tesar and Smolensky 1993,1

> Under the assumption of innate knowledge of the universal constraints, the primary task of the learner is the **determination of the dominance ranking** of these constraints which is particular to the target language. We will present a simple and efficient algorithm for solving this problem, **assuming a given set of hypothesized underlying forms**. (Concerning the problem of acquiring underlying forms, see the discussion of 'optimality in the lexicon' in P & S 1993: §9).[emphasis added—mrh&cr].

Turning to P & S 1993:§9 we find

(14) Prince and Smolensky 1993, 192

> Lexicon Optimization. Suppose that several different inputs I_1, I_2 ,..., I_n when parsed by **a grammar** G [i.e. ranked constraint hierarchy—mrh&cr] lead to corresponding outputs O_1, O_2,..., O, all of which are realized as the same phonetic form Φ — these inputs are all *phonetically identical* with respect to G. Now one

> of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labeled O_k. Then the learner should **choose for the underlying form** for **Φ,** the input I_k.

We might refer to this approach as the 'vicious circle' theory of language acquisition: the child needs a ranking to get URs and needs URs to get a ranking. To be fair, later work including Smolensky (1996) appears to address this problem, but fails for reasons we have discussed. Again, constraint reranking and choice of UR are part of the same task.

4. Summary

In this section, we summarize the major points in our argument. First, S's parsing algorithm selects only **the most 'harmonic'** UR, so it fails to account for surface ambiguity in any human language. Second, an algorithm which associates a perceived form with a set of possible URs is needed, since surface ambiguity does exist. Finally, using such an algorithm in acquisition, the learner can converge on a lexicon **only** if F-constraints are initially ranked **above** W-constraints.

One conclusion we can draw from this is that the 'emergence of the unmarked' is irrelevant to the description of children's speech. Their **grammars** are faithful to the observed target forms.

5. Conceptual problems

As we have demonstrated, S's parsing model **cannot** account for the peculiarities of child speech. However, it seems clear that such an account **should not** even be attempted in a model of grammar. We propose that the standard performance/competence distinction must be maintained, and that *contra* S (p.1), there is a "dramatically greater performance/competence gap for children than for adults." 'Performance' includes all extralinguistic cognitive and motor processing. This includes everything from motor control in articulation to motor planning and short-term memory access and allocation. See Hale and Reiss (1995, 1996abcd) for further discussion. In rejecting a performance AND competence approach, S refers to "gross formulations" of the hypothesis, but he cites no references. Instead, S refers to well-described phenomena in the child phonology literature: improved production of adult target forms during direct imitation and the existence of apparent 'chainshifts' in child speech *vis à vis* target forms. We will now argue that both these objections are irrelevant, except insofar as they support the competence and performance model and provide evidence against S's own model.

The higher level of performance during direct imitation that S cites from Menn and Matthei (1992) as further evidence for his model is actually contradictory to his own approach to the study of child speech output. There are two distinct accounts for what has been labeled 'imitation': (1) increased performance skill under concentration and (2) parroting. Under our account it is precisely during intense concentration on the act of performance that the child will perform better in carrying out the instructions provided by the grammar and thus improve articulation of the target forms. Note that since the child must have F-constraints ranked high, the target for their own grammar is identical to that of the adult, except where a performance error in parsing has led the child to posit the wrong UR. Parroting clearly has no grammatical basis: a speaker of English can parrot a Cree sentence fairly well without acquiring a Cree grammar.

Under S's competence-only approach, neither of these types of "imitation" can be accounted for. In the first type, since S assumes the *grammar* is responsible for, e.g., realization of [š] as [s], increased attention to performance should lead only to a clearer hit of the child's target [s], not to a closer approximation to the adult target [š]. In the second case, to account for an English speaker's ability to imitate Cree, S would have to assume instantaneous acquisition of Cree constraint rankings.

Smolensky also cites the following as evidence for the grammatical basis of children's speech output: children who produce 'thick' as [fɪk] cannot be said to be physically unable to produce [θ] since they produce this sound when saying 'sick' as [θɪk]. Although S cites these 'chainshifts' as problematic for proponents of non-grammatical accounts of child speech output and thus as support for his theory, he never actually demonstrates how an OT grammar can allow a child to produce, e.g. [θ] for an underlying /s/, but [f] for underlying /θ/. The treatment of chainshifts and other opacity effects has been one of the most difficult issues for OT. Reiss (1996) and others have demonstrated why a well-constrained OT grammar has difficulty with chainshifts. Simply put, the problem is this: if the optimal output for underlying /θ/ is [f], why isn't [f] also a better output for underlying /s/ than [θ] is? Or similarly, [θ] is as well-formed (with respect to W-constraints) whether it corresponds to underlying /θ/ or /s/, and it is more faithful to /θ/; therefore, it should be the optimal candidate for the realization of /θ/.[3]

Two of the best known attempts to deal with opacity in OT are McCarthy's "Remarks on Phonological Opacity on Optimality Theory" (1994) and Kirchner's work on chainshifts (1995). McCarthy adopts the use of parameterized constraints to account for opacity effects and suggests that the default state (i.e. initial state given by UG) for each constraint is one in which the parameters are set so as to minimize opacity. Kirchner's proposal also includes a radical modification to the original OT idea of a universal, innate set of constraints—constraints can be conjoined to generate complex constraints. Both of these suggestions rely on positive evidence and are treated as learned aspects of OT grammars. For example, McCarthy (1994, 6) says, "I will stipulate that the default form of a phonological constraint—the form in which it is represented in UG—has all of its level specifications set to 'surface'". Therefore, these solutions to the description of opaque systems cannot be applied to the chainshifts discussed by S since the target language provides no evidence for such shifts. In the absence of an alternative, chainshifts such as those cited by S constitute a major challenge to his own proposals. Chainshifts are not a problem for our theory because we do not require the grammar to produce the chainshifts — the child speech chainshifts are the result of performance effects. Sharon Inkelas (p.c.)

[3] Note as an aside that given a theory of phonology which contains rules which apply in an ordered derivation, chainshifts are predicted to occur. In that sense opacity has no status in a rule-based grammar. Opacity is just a point of logic, a possible result of applying rules in some order. This was recognized by Kiparsky and Menn (1977:73) who say that '[o]pacity is a property of the relation between the grammar and the data. An opaque rule is not more complex, merely harder to discover."

points out the plausibility of this suggestion, given that the set of 'chainshifts' attested in child speech are disjoint from the set found in adult phonology.

It is also worth noting in passing that both McCarthy's and Kirchner's proposals involve major adjustments to OT, whereas opacity effects in rule-based grammar are accounted for trivially. For example, McCarthy's decision to allow parameterized constraints in Optimality Theory compromises one of the signature distinctions of OT from Principles and Parameter models. The radical nature of this departure from the original OT notion of a fixed universal constraint set is not fully recognized in the literature, even by McCarthy himself, as illustrated by the following statement: "The constraints themselves are universal, except for the fixing of particular arguments within general constraint schemata; only the ranking is language particular" (McCarthy 1994, 2). Clearly, fixing of arguments *and* constraint ranking are both language particular in McCarthy's model.

Note that positing the 'marked' (in McCarthy's sense) setting of constraints to generate a chainshift in child speech requires changing the default setting of the parameters to produce a grammar that the target language provides no evidence for, in the case of the English example S cites. Becoming a competent adult speaker then requires readjusting the grammar (by resetting constraint parameters or by reranking complex constraints so low as to be inactive) so as to attain the adult grammar which has no chainshifts, just as the initial state of the grammar had none. Such 'Duke of York' models of the learning path are intuitively unappealing.

We propose, therefore, that the 'chainshifts' reported in the child speech literature arise at the level of implementation, at the interface of the phonology with other cognitive and motor systems: when the phonology provides the 'instructions' to articulate x, the implementation system may articulate y. The instruction to articulate y, in turn, may lead to the articulation of a distinct sound z. In other words, we distinguish between the output of the grammar, a linguistic representation, and the output of the organism, an acoustic signal. This distinction, between linguistic knowledge and speech behavior, forms the basis of all research in generative linguistics.

S's brief discussion of the acquisition process (1996: 12) is instructive; as with most of S's arguments, the sketch supports the approach advocated in this paper rather than S's own approach. The learning algorithm is given as follows: first, the child uses his/her ("incorrect", because all W-constraints outrank all F-constraints) grammar to parse (and produce) overt phonetic forms. Subsequently,

(15) The full structural descriptions assigned to the overt data are then used in the Error-Driven version... of the Constraint Demotion ranking algorithm (Tesar and Smolensky 1993): *whenever* the structural description which has just been assigned to the overt data (comprehension) is less harmonic than the current grammar's output (production), relevant constraints are demoted to make the comprehension parse the more harmonic. This yields a new grammar... [S 1996:12, emphasis added]

In the case discussed by S in which a child produces [ta] for the underlying representation /kæt/, but correctly parses [kæt] as /kæt/, the structural description of the production process is more harmonic (obeying highly-ranked NOCODA, *DORSAL, etc.) than the structural description of the comprehension process (which violates these W-constraints). Note, however, that with the learning algorithm given by S, such a state of knowledge ('grammar') cannot exist: *whenever* this would be the case (e.g., as soon as the child correctly parses [kæt] as corresponding to underlying /kæt/ — a necessary prerequisite to the acquisition of that lexeme — in spite of having, e.g., *DORSAL ranked high) the relevant W-constraints will be demoted below the relevant F-constraints. Thus the grammar posited by S, which produces [ta] for /kæt/ but correctly parses [kæt] as /kæt/ will never exist and cannot provide a competence-based account of stable features of child speech output of this type. Performance factors, as argued here, must be invoked. Since S's model predicts instantaneous reranking whenever a learned form is produced, it cannot account for the very data it sets out to — children's staged approximation to adult pronunciation.

6. Conclusions

We conclude that S's OT parser must be replaced by one which accounts for surface ambiguity. Furthermore, the failure of S's account undermines its claimed support for OT over other models. An interesting OT parser can be designed, but its equivalent can be formulated in other frameworks. Considering some of the OT catch-phrases in the literature we have seen that the 'emergence of the unmarked' turns out to be a misnomer. S is applying it to nonlinguistic phenomena (as are all other OT acquisition theorists (see S for references) who assume that F-constraints are ranked low initially). S cannot account for these phenomena anyway, since the child is posited to **immediately** elevate the F-constraints. In addition, the 'richness of the base' is mostly irrelevant to acquisition, and thus perhaps to mental grammar in general, under S's own assumptions, since he assumes accurate parsing and thus accurate acquisition of adult forms. The performance/competence distinction must be maintained: an explicit characterization of the boundaries between the two should be one of the primary goals of phonological theory, since it defines the sphere of inquiry with which we must concern ourselves. It is clear that a more explicit theory of performance (or rather several theories) is a necessity; however, it must be accompanied by a coherent theory of grammar which is consistent with fundamental assumptions of the field.

References

Hale, M. and C. Reiss. 1996a. Phonological Underspecification and the Subset Principle. *WECOL* 1996, UCSC.

Hale, M. and C. Reiss. 1996b. The Comprehension/Production Dilemma in Child Language: A Response to Smolensky. *ROA*.

Hale, M. and C. Reiss. 1996c. Competence and Performance in the Acquisition of Phonology. Conference on *What Children Have to Say About Linguistic Theory*. Institute of Language and Speech (OTS), Utrecht, Holland.

Hale, M. and C. Reiss. 1996d. Evidence in Phonological Acquisition: Implications for the Initial Ranking of Faithfulness Constraints. *Proceedings of the 28th CLRF.* Stanford: CSLI Publications.

Hale, Mark and Charles Reiss. 1995. On the Initial Ranking of OT Faithfulness Constraints in Universal Grammar. *ROA*.

Kiparsky, P. and Lise Menn. 1977. On the Acquisition of Phonology. In *Language, Learning and Thought*, ed. J. Macnamara. New York: Academic Press.

Kirchner, Robert. 1995. Synchronic chain shifts in Optimality Theory. *LI* 27.

McCarthy, J. 1994. Remarks on Phonological Opacity in OT. Ms. UMass, Amherst. To appear in *Proceedings of 2nd Colloquium on Afro-Asiatic Linguistics*. J. Lecarme, J. Lowenstamm and U. Schlonsky, eds.

Menn, Lise and Edward Matthei. 1992. The "two-lexicon" account of child phonology : Looking back, looking ahead. In *Phonological Development: Models, Research , Implications*, ed. Charles A. Ferguson, Lise Menn, and Carol Stoel-Gammon, 211-247. Timonium, Md.: York Press.

Prince, A. and P. Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, N.J.

Reiss, C. 1996. Deriving an Implicational Universal in a Constrained OT Grammar. *Proceedings of NELS 26*:303-17. UMass Graduate Linguistics Association, Amherst.

Smolensky, P. 1996. On the Comprehension/Production Dilemma in Child Language. To appear in *LI* 27.

Tesar, B. and P. Smolensky. 1993. The learnability of Optimality Theory: An algorithm and some basic complexity results. Technical Report, Computer Science Department, U. of Colorado, Boulder.

Department of Classics, Modern Languages and Linguistics
Concordia University
1455 de Maisonneuve West
Montreal, Quebec
H3G 1N1
Canada

hale1@alcor.concordia.ca, reiss@alcor.concordia.ca