JOHNS HOPKINS
UNIVERSITY

Department of Cognitive Science

☎ (410) 516-5250 Baltimore, MD 21218-2685 Fax: (410) 516-8020

The Initial State and 'Richness of the Base' in Optimality Theory

Paul Smolensky

October 1996

Technical Report
JHU-CogSci-96-4

http://www.cogsci.jhu.edu/TechReports/

TR-Request@cogsci.jhu.edu

The Initial State and 'Richness of the Base' in Optimality Theory

Paul Smolensky

Basic learnability considerations are argued to explain the broad generalization that the same linguistic structures which are marked in the sense of later-emerging in child language tend also to be marked in adult languages. Using Optimality Theory (OT), and developing a proposal of Prince (1993), this generalization can be reduced to the requirement that the initial state $\mathcal{H}_0$ possess the property that structural markedness constraints outrank faithfulness constraints. $\mathcal{H}_0$ is explained as a learnability consequence of a fundamental OT principle, *richness of the base*: the set of possible inputs to the grammar is universal. This principle entails a strong requirement for what it means to have acquired a language with an unmarked inventory: subordination of faithfulness constraints in the absence of evidence of their domination. This raises a familiar type of Subset Problem for acquisition, which $\mathcal{H}_0$ resolves. Richness of the base lends unity to a seemingly incoherent set of assumptions defining the emerging OT theory of acquisition: the initial state is $\mathcal{H}_0$, and the child's inputs to the grammar are close to the adult form.

Keywords: learnability, markedness, Optimality Theory, Subset Problem, richness of the base, acquisition

A fundamental link between child and adult language is provided by what I will call the *Jakobsonian Generalization*: broadly speaking, the same structures are 'marked' in adult and child grammars. That is, structures which are avoided via phonological "processes" within adult languages, and excluded from some inventories across adult languages, also tend to be structures which are later-acquired by children (Jakobson 1941/1968, Stampe 1979). This generalization is of course a most broad one, one which abstracts away from many interesting phenomena within child language. My topic is not an assessment of the empirical scope and limits of the generalization, but rather the question of whether it can be explained by comparably broad, but more fundamental, theoretical considerations.

In many respects a formal theory of markedness, Optimality Theory ('OT'; Prince and Smolensky 1991, 1993) is a reasonable place to look for such an explanation. Indeed, current OT acquisition research (Demuth, in press, Pater and Paradis 1996, Bernhardt and Stemberger 1995, Gnanadesikan 1995, Levelt 1995) in essence provides a straightforward formal expression of the Jakobsonian Generalization, as follows. Child and adult grammars consist of the same (minimally violable) markedness-defining *structural constraints*: these are indeed universal constraints, the core of UG. The initial state is characterized by a special property: these structural constraints outrank constraints requiring *faithfulness* between the inputs and outputs of the grammar. Throughout this paper, I will use $\mathcal{H}_0$ to denote a constraint hierarchy with this property.[1]

Because child grammars begin as $\mathcal{H}_0$, faithfulness to inputs cannot force early child outputs to violate the structural constraints; grammatical outputs cannot incur structural constraint violations or *marks*: in this precise sense, they are *unmarked* structures.[2] Thus early child outputs avoid the same structural constraint violations as those of adult languages in which structural constraints are sufficiently highly ranked: this is the Jakobsonian Generalization.

During acquisition, the child's constraint ranking changes to match the adults', and as structural constraints are demoted below faithfulness constraints, those marked structures which may appear in the target adult language also emerge in the child's outputs. (Learning procedures for performing this reranking are briefly discussed below.) Those marked structures which are absent in the target language are avoided in the child's ultimate grammar too, since the structural constraints defining those marked structures end up more highly ranked than faithfulness constraints, just as they are in $\mathcal{H}_0$.

Thus the Jakobsonian Generalization can be expressed succinctly within OT as the following principle: the initial state is $\mathcal{H}_0$. To explain the Generalization, then, we must explain this principle: such explanation is the subject of this paper.

Aside from its role in capturing the broad Jakobsonian Generalization, the assumption that the initial state is $\mathcal{H}_0$ is proving fruitful for the more detailed empirical study of child phonology; it is a standard assumption in the OT acquisition literature cited above. The hypothesis that learning is constraint reranking is also proving fruitful in such empirical studies; in addition, it has proved central to the development of learning algorithms and formal learnability results mentioned below.

Richness of the base. The question to be addressed here is whether the initial ranking $\mathcal{H}_0$ can be explained by other principles. Following a proposal by Prince (1993), I will develop a learnability argument showing that the initial ranking must be $\mathcal{H}_0$, or certain languages would not be learnable. This argument depends on a fundamental principle of OT which governs an aspect of grammar whose relevance to the initial state may not at first be apparent: the *inputs* to the grammar. This principle was proposed in Prince and Smolensky 1993:191; the formulation I adopt here is given in (1).

(1) **Richness of the Base.** The source of all systematic cross-linguistic variation is constraint reranking. In particular, the set of *inputs* to the grammars of all languages is the same. The grammatical inventories of a language are the *outputs* which emerge from the grammar when it is fed the universal set of all possible inputs.

Richness of the base requires that systematic differences in inventories arise from different constraint rankings, not different inputs. The lexicon of a language is a sample from its inventory: all systematic properties of the lexicon thus arise indirectly from the grammar, which delimits the inventory from which the lexicon is drawn. There are no independent morpheme structure constraints on phonological inputs; no independent lexical parameter which determines whether a language has *pro*. Apparent cross-linguistic differences in inputs are actually logical consequences of differences in constraint ranking.

A primary motivation for richness of the base is theoretical restrictiveness. In OT, languages differ in the rankings of the universal constraints that define their grammars. The most restrictive theory limits cross-linguistic variation to this one locus: constraint ranking.

For example, variation in the phonemic inventory is derived in OT from constraint reranking (Prince and Smolensky, 1993, Chapter 9): it is not a dimension of variation that is independent of grammatical reranking. Thus, variation in whether an inventory contains voiced obstruents is governed by the relative ranking of two (independently required) constraints: a structural constraint on feature co-occurrence (e.g., [voice] ⊃ [sonorant]), and a faithfulness constraint violated when features in the input are not realized in the output (e.g., PARSE$^{\text{Feat}}$). In all languages, input feature combinations are available which, if faithfully parsed, would surface as voiced obstruents; but if the faithfulness constraint is dominated by the structural constraint, such faithful parses will be sub-optimal relative to unvoiced outputs, which meet the structural constraint while sacrificing lower-ranked faithfulness. Thus the typology of segmental inventories explained by reranking includes (a) inventories with both voiced and unvoiced obstruents (faithfulness ≫ structural markedness), and (b) inventories with only unvoiced obstruents (structural markedness ≫ faithfulness) — but no inventories with only voiced obstruents: no ranking yields such an inventory. This is OT's markedness link between a violable markedness constraint like [voice] ⊃ [sonorant] and the descriptive implicational universal for inventories, "voiced obstruents ⊃ unvoiced obstruents": for other phonological examples, see Prince and Smolensky 1993, Chapter 9 (and Kirchner 1995). As Prince and Smolensky also show in that chapter, optimality principles operating in the *lexicon* entail the following result: if the grammar yields an inventory with only unvoiced obstruents, no segments *in lexical forms* will contain [voice] without [sonorant] — even though all feature combinations are universally available as inputs. (See also Itô, Mester and Padgett 1995 for discussion of both lexicon optimization and the converse structural constraint, [sonorant] ⊃ [voice].)

Richness of the base fully generalizes this conception of the roles of grammar and input: *all* systematic cross-linguistic variation is the result of reranking of universal constraints, the inputs to the grammars of all languages being the same. (A universal theory of the input is thus at least implicit in OT analyses.)

Two examples from syntax may help to underscore the generality of this principle; in this context I can only offer an oversimplied sketch of one small bit of each analysis, of course.

According to Grimshaw and Samek-Lodovici (in press, 1995; Samek-Lodovici 1995), languages with and without null subjects differ only in the ranking of a common set of syntactic constraints: the inputs to the grammar are the same, and there is no parametric contrast involving *pro*, or the ability of functional heads to govern, that is independent of ranking. One key structural constraint in this analysis prohibits topic-referring thematic subjects from being overtly expressed; this conflicts with other constraints, one that requires clauses to have overt subjects (a descendant of the EPP) and another, a type of faithfulness constraint analogous to a phonological constraint against deletion, that requires overt expression of input predicates and arguments. The relative ranking of these constraints determines whether certain structures lacking overt subjects are optimal and hence grammatical; this is part of Grimshaw and Samek-Lodovici's overall analysis of subjects.

In the same explanatory vein, the *wh*-chain theory of Legendre, Smolensky, and Wilson (in press; Legendre et al. 1995) addresses the contrast between languages with and without resumptive pronouns (more specifically, overt traces) These do not differ on some independent lexical parameter — only the ranking of constraints governing the distribution of empty traces differ. Here, a structural constraint (a version of the ECP) prohibits ungoverned empty traces, while a type of faithfulness constraint (analogous to a phonological constraint against epenthesis) prohibits the kind of double overt realization represented by these resumptive pronouns. The grammar determines what elements surface (and where); the lexicon must follow the grammar, providing lexical entries for the elements appearing in grammatical structures. Depending on the ranking, these elements may or may not include, e.g., overt traces.

Outline of the argument. Another argument for richness of the base emerges from the learning considerations which, as I now show, link this principle to the explanation of the initial ranking $\mathcal{H}_0$. In the remainder of the paper, I first develop a consequence of richness of the base for what it means to have acquired a language with an unmarked inventory. This is then shown to raise a rather familiar Subset Problem for acquisition. The assumption that the initial state is $\mathcal{H}_0$ is then shown to resolve this problem. A summary of the emerging OT theory of acquisition concludes the paper, with a focus on how richness of the base lends unity to a seemingly incoherent set of assumptions.

The argument to be presented develops a proposal of Alan Prince (1993):

Unranked initial state and 'Richness of the Base'.

A[n] … important assumption [in the Tesar and Smolensky 1993 learnability work] is that the learner actually sees the crucial evidence.

Consider the following situation. The lexicon of language L contains only morphemes constructed from {CV}*. I.e. stems are CV, CVCV, CVCVCV, suffixes are CV, CVCV, etc. Nothing prevents there from being such a language.

Every grammar parses this language, because every grammar parses CV.

Question: which grammar does the learner actually learn?

One answer might be: any grammar. … But it is doubtful whether this is correct. Learners exposed to such a language do not know how to handle closed syllables or onsetless syllables; evidence would include the fate of borrowings into the language (readily available), experimental study of native speakers, etc. Thus, it is plausible to suppose that they have learned only the narrowest grammar, the one that *only* admits .CV. syllables. (One might also be able to determine empirically which Faithfulness constraints they prefer to violate.)

If this is right, then there must be an initial state. It might look like this: STRUCTURE ≫ FAITHFULNESS, where STRUCTURE is the set of constraints like ONSET and NOCODA that delimit favored structural options.

When has a language been acquired? Acquiring a language certainly requires that inputs from the lexicon be assigned their correct structural descriptions by the learned grammar. But is this sufficient? According to richness of the base, the answer is *no*.

To see this, consider a simple example in the acquisition of C/V syllable structure; I will assume the OT analysis of Prince and Smolensky (1993:ch. 6). In this example, we abstract away from all segmental phonology, and consider only the syllabification of abstract consonants and vowels. Consider a language we will call $\Sigma^{/CV/}$, a language in which all syllables have the unmarked form .CV. (i.e., $[_\sigma CV]$), and in which there are no alternations: no segments are "deleted" or "epenthesized."[3] In such a language, there is no need to posit deep/surface disparities: the underlying form of morphemes can be identical to their surface form.

When can a learner be said to have acquired this language $\Sigma^{/CV/}$? The learner's grammar must certainly take an input such as /CVCV/ and assign it the correct structural description, .CV.CV. But is this sufficient? Note that in $\Sigma^{/CV/}$, *all* lexical items are of the form /CVCV··· CV/, so the criterion of correct parsing of lexical items is a particularly weak one. Indeed, as observed by Prince and Smolensky, *any* ranking of the syllable structure constraints will meet this criterion: the correct outputs are faithful to the inputs, violating no faithfulness constraints, and the output syllables meet the structural constraints barring codas and requiring onsets — they are structurally unmarked. For the lexical items of $\Sigma^{/CV/}$, the correct outputs violate no constraints, and are therefore optimal under all constraint rankings.

Is this conclusion correct, that *any* ranking at all counts as having acquired the language $\Sigma^{/CV/}$? No, according to richness of the base. In this language, syllables are characterized by a strong regularity: they all have the unmarked form, .CV. Richness of the base (1) demands that this regularity result from the grammar alone, assuming no limitations on inputs. An arbitrary ranking can output just .CV. syllables — but only when the inputs are themselves strongly restricted, to the form /CVCV··· CV/. To produce just .CV. syllables with *unrestricted* inputs, as required by richness of the base, the grammar of $\Sigma^{/CV/}$ must obey a strong restriction: the syllable-structure constraints must outrank faithfulness constraints.[4] Encapsulating the structural constraints determining structural markedness (or, in OT terminology, *structural Harmony*) under the name STRUC-H, and encapsulating the faithfulness constraints as FAITHFULNESS, we have the following schematic ranking:

(2) $\Sigma^{/CV/}$ learned only under the ranking: STRUC-H ≫ FAITHFULNESS

Under such a ranking, given an *arbitrary* input (like /CVCVC/), structural constraints will force unfaithful parsing as needed to yield only .CV.-syllables.[5]

Here we see one reflex of the pressure for theoretical restrictiveness that motivated richness of the base in the first place. Since language-particular restrictions on the inputs cannot be appealed to, only language-particular *ranking* can meet the criterion of 'having learned' the language $\Sigma^{/CV/}$. And this will now enable us to draw a conclusion about what initial ranking will allow such languages to be learned.

The problem of learning unmarked inventories. It is clear that meeting the learning criterion entailed by richness of the base is a challenge. The learner of $\Sigma^{/CV/}$ has a lexicon consisting only of forms like /CVCV··· CV/. There are no forms like /CVCVC/ which could provide evidence — by surfacing unfaithfully (e.g., as .CV.CV.⟨C⟩) — for the necessary ranking (2), in which syllable structure constraints dominate faithfulness. With no morphologically-induced alternations such as that shown in (3), there are no faithfulness violations with lexical items, thus no evidence that FAITHFULNESS is dominated.

(3) Typical alternation motivating violation of FAITHFULNESS in optimal forms:

/CVCV**C**+V/ → .CV.CV.**C**+V.

/CVCVC/ → .CV.CV.⟨**C**⟩

This learning problem would not arise with a reversed target ranking, in which FAITHFULNESS dominates some structural constraints. For example, if the inventory of syllables included marked syllables like .CVC., these would provide evidence for the correct ranking: a structural constraint NOCODA is violated in an optimal form, and it may therefore be deduced that this constraint must be dominated by a faithfulness constraint which forces its violation. But in $\Sigma^{/CV/}$, with the inventory of syllables limited to the *unmarked* syllable .CV., there are no surface violations of STRUC-H; if alternations like (3) do not happen to provide the necessary evidence, as we have assumed they do not in our language $\Sigma^{/CV/}$, then we have no evidence from which to deduce the correct ranking. Clearly, this constitutes a kind of Subset Problem for learnability (Angluin 1978, Berwick 1986, Pinker 1986, Wexler and Manzini 1987). Because the inventory contains only the unmarked structure, it is possible that there are no constraint violations or 'marks' in the positive data, and therefore no evidence for the target constraint ranking (2) required by richness of the base.

To make this problem even more concrete, suppose a learner of the all-.CV. language $\Sigma^{/CV/}$ follows an error-driven learning procedure (such as the Error-Driven Constraint Demotion algorithm of Tesar, in press). Suppose our learner has acquired the underlying forms of some words: not too challenging, since, with respect to C/V structure, these are identical to their surface form. As we have seen, these underlying forms will consist entirely of CV-sequences (e.g., /CVCV/), and regardless of the learner's ranking, all these inputs will be correctly parsed (e.g., .CV.CV.). Thus the learner will make no errors, and hence no learning (reranking) can occur. Regardless of the quantity of positive language data provided, the learner will end up with the same ranking as she started with.

Thus the only way the learner can end up with a correct ranking is if she already had one to start with. That is, the language $\Sigma^{/CV/}$ can be learned only if the *initial* ranking satisfies the requirement (2). Thus:

(4) Initial state $\mathcal{H}_0$: STRUC-H ≫ FAITHFULNESS

Loanwords. For another perspective on this conclusion, suppose the initial state were a ranking opposite to (4): suppose the FAITHFULNESS constraints top-ranked. This then would be true of the $\Sigma^{/CV/}$-learner's final grammar too, since no reranking occurs. What would happen if a new word entered the language from another language in which codas are possible? If, say, /CVCVC/ were adopted as the underlying form of a loan word originally pronounced .CV.CVC., then the high-ranking FAITHFULNESS constraints would ensure that this form surfaces as .CV.CVC. Indeed, all aspects of the loan word would be taken over with no adaptation whatever to the constraints of the receiving language, since FAITHFULNESS is presumed top-ranked. Clearly this is the wrong result: while loanwords do not necessarily respect *all* the constraints of the receiving language, there is no doubt that the correct generalization is that they are strongly reshaped by those constraints (see Paradis 1995 and Yip 1993 for recent constraint- and optimality-based discussions).

At this point we digress to develop, and reject, one plausible response to this problem. It might be suggested that the error here is due not to assuming an initial ranking with FAITHFULNESS top-ranked, but rather to the assumption that a speaker of an entirely-.CV. language, when borrowing a word with foreign surface shape .CV.CVC., would take the underlying form to be /CVCVC/. This underlying form is certainly a reasonable possibility, given that in the native vocabulary there are in fact no deep/surface disparities. On the other hand, it is also true that in the native vocabulary all underlying forms are CV sequences, respecting a constraint which is violated by /CVCVC/. So perhaps the learner would acquire a morpheme structure constraint which prevents an underlying form /CVCVC/ from being adopted. That is, perhaps some constraint we can call 'C ⊃ CV' ('C only if following V') would apply to underlying forms. But in borrowing a word with shape .CV.CVC., this constraint would clearly conflict with a fundamental constraint on deriving underlying forms, a constraint which might be dubbed NO-DEEP-SURFACE-DISPARITY. Now the alternative under development here would require that the selected underlying form be /CVCV/, and to get this result, we would have to assume that C ⊃ CV has priority over NO-DEEP-SURFACE-DISPARITY. In developing this little theory of how underlying forms would be selected, it is now clear that we would be heading down the road to duplicating the whole machinery of the grammar: on the formal side, reconstructing constraint ranking to resolve conflict; on the substantive side, duplicating the grammatical constraint NOCODA with the morpheme structure constraint C ⊃ CV, and duplicating FAITHFULNESS constraints with NO-DEEP-SURFACE-DISPARITY. The result we seek — broad assimilation of loanwords to native constraints — would arise only if the structural-type constraint (C ⊃ CV) were to out-rank the faithfulness-type constraint (NO-DEEP-SURFACE-DISPARITY). Thus, even after all this duplication of the grammatical apparatus — raising concerns similar to those identified with the 'duplication problem' of Kenstowicz and Kisseberth (1977:136–149), a major target of Kiparsky (1982) and subsequent work in the development of lexical phonology — even after large-scale grammatical duplication, in the end, we would end up having to stipulate essentially what we set out to explain in the first place: why structural constraints have priority over faithfulness constraints in the absence of learning data concerning how their conflicts are resolved.

In theoretical phonology under OT, a major role of richness of the base is to do the work of morpheme structure constraints. What the considerations of the previous paragraphs show is that, in the context of learning $\Sigma^{/CV/}$, the requirement deriving from richness of the base, (2), perspicuously yields the broad loanword generalization without necessitating a highly redundant theory of the acquisition of morpheme structure constraints.

But the possibility of lack of ranking evidence in languages like $\Sigma^{/CV/}$ with unmarked inventories entails that the demands of richness of the base can be met only if the initial state has the special structure of (4).

The special status of FAITHFULNESS. Why do FAITHFULNESS constraints merit special consideration in the initial ranking? As we have seen, if STRUC-H is low-ranked in the target grammar, evidence to this effect will be available; on the other hand, if FAITHFULNESS constraints are low-ranked in the target grammar, evidence to this effect may be unavailable, so these constraints must start out low-ranked. What is the basis of this asymmetry?

As manifest both in linguistic analysis and in the development of learning algorithms within OT, evidence for constraint ranking takes the following form. Positive data from the target language provides the optimal parse of some input. This may be compared with a competing suboptimal parse. Each (uncancelled) constraint violation incurred by the optimal parse must be lower ranked than some (uncancelled) constraint violation incurred by the suboptimal parse (Prince and Smolensky 1993:221; Tesar and Smolensky 1993:10). The key question is, what competitors are available for this comparison?

The OT generator of output candidates, *Gen*, displays what McCarthy and Prince (1993) dub 'freedom of analysis': given any input, *Gen* provides a wide range of competitors violating the various structural constraints of UG. If a structural constraint like NOCODA is low ranked in the target grammar (relative to FAITHFULNESS), closed syllables like .CVC. will be optimal, marked though they are by the violation of NOCODA. *Gen* provides alternative parses that avoid this mark (via unfaithful parsing), so the positive evidence can always be compared with *Gen*-provided alternatives that reveal the low-ranking of NOCODA.

The situation is different with FAITHFULNESS constraints. If such a constraint is low-ranked, that may or may not entail that optimal forms will violate the constraint. If alternations like (3) are present, as we have seen, some outputs will indeed display FAITHFULNESS violations, and these can be compared to *Gen*-provided alternatives which lack FAITHFULNESS violations, yielding the necessary ranking evidence. But the existence of this evidence cannot be guaranteed by *Gen*'s freedom of analysis: what is needed is the presence in the lexicon of inputs like /CVCVC/, input which cannot be faithfully parsed with unmarked structure. Yet such inputs may be absent from the lexicon, as in $\Sigma^{/CV/}$ — a systematic gap arising indirectly from the language's constraint ranking, when morphology does not lead to alternations. It is the characteristic property of FAITHFULNESS constraints that they crucially inspect the *input*; thus it is the FAITHFULNESS constraints that are the target of the learning difficulties that arise from an insufficiently rich input base in the lexicon.[6]

Learnability and the initial state in Principles-and-Parameters Theory and OT. The Subset Problem identified here, and its resolution via an initial state that yields inventories of only unmarked structures — subsets of richer possible inventories — is clearly related to proposals in the learnability literature within the Principles-and-Parameters framework (e.g., Berwick 1986, Pinker 1986). While the relation between the OT and P&P accounts is more complex than may at first be apparent, I will confine myself to a few remarks.

The OT initial state $\mathcal{H}_0$ is based on a distinction which is fundamental to the theory, the distinction between structural and faithfulness constraints. The condition that the former dominate the latter requires no special considerations or apparatus not already needed in the theory. In contrast, in the standard proposal of P&P theory, initial parameter values can be determined only when there is a clear subset relation between the languages generated by alternative parameter values. Such subset relations are by no means guaranteed by the limited structure provided by the P&P framework; this creates both conceptual and technical challenges for learnability theory (Frank and Kapur, 1996). One response is to elevate to the status of a meta-principle of P&P theory the requirement that parameters, independently of one another, *must* give rise to superset/subset languages (Manzini and Wexler 1987). This meta-principle seems at odds with actual P&P proposals, at least in part because it typically conflicts with a main goal of linguistic theory: formulation of cross-cutting, interacting principles which unify diverse surface phenomena (see Tesar and Smolensky 1996, forthcoming).

In contrast, interactions of constraints in OT is the heart of both linguistic theory *and* learning: the learning algorithms of Tesar and Smolensky 1993, 1996 and Tesar, in press, are based entirely on constraint interaction. As discussed above, these procedures converge on adult grammars by demoting constraints in the face of evidence of constraint interaction (ranking) different from that of the currently hypothesized grammar. The new element proposed here is the requirement that the initial ranking be $\mathcal{H}_0$. This new requirement can be added to the existing theory of OT learning algorithms without introducing new apparatus or difficulties, conceptual or technical. (Previous work on Constraint Demotion algorithms assumed the initial ranking to be one in which all constraints are equally ranked. However, the type of formal learnability results derived in that work can be extended from that special initial ranking to the general case of an arbitrary initial ranking of the universal constraints: see Tesar and Smolensky 1996, forthcoming).[7]

Finally, the difference between the use of ‘unmarked’ in OT and P&P theory should be noted. In P&P, the notion of an ‘unmarked’ parameter value as one yielding a subset language is a construct external to the operation of the grammar itself. In OT, the structures surfacing under the initial grammar are structurally ‘unmarked’ in a sense which is fundamental to the theory: these structures best satisfy the structural constraints, according to the formal evaluation procedure which constitutes the OT grammar itself.

Learning theory in OT. In Optimality Theory, learning a target adult language requires a child to determine the relative rankings of universal constraints. When a faithfulness constraint outranks a structural constraint in the target grammar, positive evidence for this ranking will appear in the form of grammatical structures violating the structural constraint: marked structures will appear. However, when structural constraints STRUC-H outrank faithfulness constraints, marked structures do not surface, and positive evidence for this ranking may be lacking, in the absence of alternations which entail surface violations of FAITHFULNESS. Learnability thus requires the child’s initial hierarchy to rank FAITHFULNESS below structural constraints.

With this initial ranking, child productions consist of unmarked structures, that is, structures best-satisfying the constraints STRUC-H. These are the same structural constraints which govern adult language. In some cases, a structural constraint $\mathbb{S}$ (e.g., NOCODA) will also outrank a faithfulness constraint $\mathbb{F}$ in the *target* language: in this case, the adult language bars from its inventory the same $\mathbb{S}$-violating structures as child language (e.g, closed syllables). The same structures which are 'marked' in the sense of barred from certain adult inventories are also the structures absent from early child inventories.

On the other hand, a structural constraint $\mathbb{S}$ may be *lower* ranked in the target language than a faithfulness constraint $\mathbb{F}$. In this case, the structures marked in virtue of violating $\mathbb{S}$ are not banned in the adult language — but, often, the distribution of these marked elements will be restricted; they can appear only in those environments where $\mathbb{S}$ cannot be satisfied without violating higher-ranked constraints such as $\mathbb{F}$. In traditional terms, "phonological processes" prevent this marked structure from appearing outside certain environments. As the child learns such a target ranking, $\mathbb{S}$ is demoted below $\mathbb{F}$, and the marked structure then emerges in the child's productions. The same structures which are 'marked' in the sense of avoided by phonological "processes" are also the structures which emerge later in child language.

This account of child productions thus explains the Jakobsonian Generalization. What about the *inputs* to early child grammars? The low relative ranking of faithfulness in the initial ranking causes child productions to be quite unfaithful to their inputs, and would seem also to entail that the child inputs must also be quite unfaithful renderings of their adult counterparts. This is not the case, however. In Smolensky 1996, it is shown how, even under the initial grammar, during *comprehension* of adult surface forms, the parse assigned by the child's grammar is a highly faithful one. The strong comprehension/production disparity regarding faithfulness to the adult form results from the difference in *competitor sets* in the two cases; there is only one child grammar, only one ranking. This analysis allows OT not only to render coherent, but in fact to derive from more basic principles, the seemingly implausible assumption — central to much of the ongoing OT research in the acquisition of phonology (Demuth, in press, Pater and Paradis 1996, Bernhardt and Stemberger 1995, Gnanadesikan 1995, Levelt 1995; see also Smith, 1973) — that early child inputs are quite faithful to the adult forms, despite the dramatically unfaithful character of their outputs.

Combining this result with those of the present paper, we see that it is now possible to explain from more fundamental principles the two basic assumptions underlying most current OT acquisition work:

(5) Basic assumptions of OT acquisition theory

a. The initial ranking is $\mathcal{H}_0$: STRUC-H ≫ FAITHFULNESS.

b. Child inputs are faithful to adult surface forms.

As we have seen, (5a) captures the Jakobsonian Generalization, and is in turn explained by richness of the base. And this pair of assumptions is also seen to be consistent with richness of the base in placing the burden of explaining the systematic deviations between child productions and adult targets on special *ranking*, not special *inputs*: the child's inputs are generally close to the adults' — it is a different *grammar* that is responsible for explaining the systematic deviations of the child's language.

The basic acquisition assumptions (5) together generate the final conceptual puzzle to be addressed here. If children already know the correct adult form of a word, as the lexical input (5b), why can't they just promote FAITHFULNESS to the top of the ranking (5a)? With FAITHFULNESS top-ranked, their outputs would then be faithful to their inputs, which in turn are "faithful" to adult forms; thus, they would trivially solve their learning problem, and immediately produce correct adult forms.

But would they, indeed, solve their learning problem this way? The child's real job, of course, is not to learn to imitate adult productions of a given stock of words, but to *learn the target grammar*. And this problem is a difficult one, not solvable by simply bumping FAITHFULNESS to the top of the ranking. The adult grammar richly interleaves the family of faithfulness constraints among various structural constraints: the target hierarchy yields a mixture of marked and unmarked inventories along numerous structural dimensions, with the distributions of marked elements regulated in subtle ways by the exact ranking.

And this fact, that adult grammars do not simply rank all faithfulness constraints at the top, but interleave them among structural constraints, is presumably a fact about the cognitive role of grammars in the first place. Inventories of entirely unmarked structures might be most easily processed in production or comprehension, but they do not allow for the range of distinctions required for the expressiveness of adult language; functional considerations thus prevent faithfulness constraints from all being lowest-ranked in the adult grammar.

On the other hand, placing all faithfulness constraints at the top of the hierarchy would lead to completely unrestricted inventories, in which the full universal range of input distinctions are faithfully rendered on the surface. The result would be a language in which structural constraints have minimal rein on outputs, a language in which almost anything goes: all universally possible phonemic constrasts, all universally possible syllable structures, and so forth, would be faithfully expressed in surface forms. There are good reasons to believe that a lexicon dispersed through such an enormous space of forms would not be a lexicon the human memory system could store. The difficulty of storing unsystematically related items in human memory is a factor invoked, for example, in explaining why presumably memorized lexical exceptions — e.g., to the regular English past tense inflection — tend to be grouped into similarity-based clusters, and why successful storage of exceptional items requires relatively high item frequency (see, e.g., Pinker and Prince 1993; for possible neural network explanations, see Rumelhart and McClelland 1988 and in a different vein Brousse and Smolensky 1989). Thus it seems a relatively safe speculation that the requirement that lexical underlying forms be storable in human memory provides significant cognitive pressure to minimize the number of contrasts which surface in any given language.

Fortunately, thanks to combinatorial explosion, generating the range of distinctions necessary for adult expressiveness does not require ranking *all* structural constraints below faithfulness constraints; languages generate quite adequate ranges of surface distinctions by ranking only a subset of the structural constraints below faithfulness constraints. This yields surface forms marked to higher degrees on certain structural dimensions, lower degrees on others. Such compromises between excessive cognitive load and inadequate expressiveness can be struck in myriad ways: these are the constraint rankings functionally suitable for mental life in the adult world. The child's job is to determine which of these compromises has been adopted in her target language. This she must do, I have argued, by starting with structural constraints above faithfulness constraints, and demoting structural constraints only as needed to admit into her inventory those unmarked structures evidenced in the target adult language.[8]

Notes:

Thanks to Adamantios Gafos, Géraldine Legendre, John McCarthy, and Bruce Tesar for helpful conversations and suggestions, and to Linda Lombardi for useful comments on the manuscript. Special thanks are due Alan Prince for his generosity with the key idea and for most helpful suggestions on the manuscript; credit for insight should be assigned to Alan, blame for errors to me. For partial financial support, I gratefully acknowledge the Center for Language and Speech Processing at Johns Hopkins, NSF grant BS-9209265, and subsequently NSF grant IRI-9213894.

1. The property defining $\mathcal{H}_0$, that structural constraints dominate faithfulness constraints, is of course highly schematic; ultimately, a considerably more articulated formulation of this property may be anticipated. Given the broad character of the Jakobsonian Generalization which is the target of this paper, however, the schematic form will suffice.

2. On the OT formalization of unmarkedness as lack of marks, see Smolensky 1993. The basic claim is that whereas the underspecification theory of unmarkedness achieves the relative "invisibility" of unmarked structure to phonological "processes" by stipulating that unmarked structure is simply absent from representations (until filled in), in OT it *follows* from the basic operation of the grammar that structure which violates no constraints — receives no marks — is invisible to the grammar: evaluation of optimality is based solely on marks, so structure that receives no marks is literally invisible to the grammar's operation.

3. John McCarthy (p.c. 1995) suggests that Kikuyu provides a good approximation.

4. More precisely, each of the basic syllable structure constraints must outrank one of the faithfulness constraints; see Prince and Smolensky 1993:sec. 6.2.2.

5. The output will be either .CV.CV.〈C〉, with an unparsed/"deleted" final C, or .CV.CV.C□́., with an empty/"epenthesized" final nucleus — depending on details of how faithfulness constraints are ranked relative to one another; see Prince and Smolensky 1993, Chapter 6.

6. If UG contains constraints other than FAITHFULNESS which refer specifically to the distinction between input material and material supplied by *Gen*, the learnability arguments presented here may be extensible to them as well.

7. When started with any initial ranking, the Constraint Demotion algorithms converge to a ranking that correctly accounts for the optimality of all available positive learning data. The number of informative examples needed for convergence is at worst twice that required by an initial ranking in which all constraints are equally-ranked: at most $N(N\text{–}1)$ such examples, where N = number of constraints.

8. With respect to a rather abstract measure of distance between constraint hierarchies, Constraint Demotion converges monotonically to a correct hierarchy: the distance steadily decreases. Nonetheless, Constraint Demotion can produce a quite complex course of acquisition, as the relative ranking of a particular pair of constraints, for example, can flip back and forth. Thus it does *not* follow from this general learning theory that the child's inventory will strictly monotonically increase, with more and more marked items entering the inventory, and none leaving.

References

Angluin, Dana. 1978. Inductive inference of formal languages from positive data. *Information and Control* 45:117–135.

Bernhardt, Barbara H., and Joseph P. Stemberger. 1995. *Nonlinear phonology and phonological development: A constraint-based approach.* Ms., University of British Columbia, Vancouver and University of Minnesota, Minneapolis, Minn.

Berwick, Robert. 1986. *The acquisition of syntactic knowledge.* MIT Press, Cambridge, MA.

Brousse, Olivier, and Paul Smolensky. 1989. Virtual memories and massive generalization in connectionist combinatorial learning. *Proceedings of the Eleventh Annual Meeting of the Cognitive Science Society*. 380–387.

Demuth, Katherine. In press. Markedness and the development of prosodic structure. In *Proceedings of NELS 25*, 13–25. GLSA, University of Massachusetts, Amherst.

Frank, Robert, and Shyam Kapur. 1996. On the use of triggers in parameter setting. *Linguistic Inquiry*, 27.

Grimshaw, Jane, and Vieri Samek-Lodovici. 1995. Optimal subjects. In J. Beckman, S. Urbanczyk, and L. Walsh, eds., *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, GLSA, University of Massachusetts, 589–605.

Grimshaw, Jane and Vieri Samek-Lodovici. In press. Optimal subjects and subject universals. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, eds., *Is the Best Good Enough? Proceedings of the Workshop on Optimality in Syntax*. MIT Press and MIT Working Papers in Linguistics.

Gnanadesikan, Amalia. 1995. Markedness and faithfulness constraints in child phonology. Ms., University of Massachusetts, Amherst.

Itô, Junko, R. Armin Mester, and Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26, 571–613.

Jakobson, Roman. 1941/1968. *Child language, aphasia and phonological universals*. The Hague: Mouton.

Kenstowicz, Michael and Charles Kisseberth.1977. *Topics in Phonological Theory*. New York: Academic Press.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In I.-S. Yang, ed., *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co.

Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer, and William Raymond. 1995. Optimality in *wh*-chains. In *University of Massachusetts occasional papers in linguistics: Papers in Optimality Theory* 18, 607–636. GLSA, University of Massachusetts, Amherst .

Legendre,Géraldine, Paul Smolensky, and Colin Wilson. In press. When is less more? Faithfulness and minimal links in *wh*-chains. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, eds., *Is the Best Good Enough? Proceedings of the Workshop on Optimality in Syntax.* MIT Press and MIT Working Papers in Linguistics.

Levelt, Clara. 1995. Unfaithful kids: Place of Articulation patterns in early child language. Paper presented at the Department of Cognitive Science, Johns Hopkins University, Baltimore, Md., September, 1995.

McCarthy, John, and Alan Prince. 1993. *Prosodic Morphology I: constraint interaction and satisfaction*. Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ. To appear as Linguistic Inquiry Monograph, MIT Press, Cambridge, MA.

Paradis, Carole. 1995. Derivational constraints in phonology: Evidence from loanwords and implications. *Proceedings of the 31st Annual Meeting of the Chicago Linguistic Society.*

Pater, Joe, and Johanne Paradis. 1996. Truncation without templates in child phonology. In *Proceedings of the Boston University Conference on Language Development* 20, 540–552. Somerville, MA: Cascadilla Press.

Pinker, Steven. 1986. Productivity and conservatism in language acquisition. In *Language learning and concept acquisition,* ed. W. Demopoulos and A. Marras, Ablex, Norwood, NJ.

Pinker, Steven, and Alan Prince. 1993. Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society*. 230–251.

Prince, Alan. 1993. Internet communication, September 26.

Prince, Alan, and Paul Smolensky. 1991. Notes on connectionism and Harmony Theory in linguistics. Technical Report, Computer Science Department, University of Colorado at Boulder. [Course notes, LSA Linguistic Institute, UC Santa Cruz.]

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Technical Report, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, N.J., and Computer Science Department, University of Colorado, Boulder.

Rumelhart, David E. and James L. McClelland. 1986. On learning the past tenses of English verbs. In James L. McClelland, David E. Rumelhart, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*, 216–271. Cambridge, MA: MIT Press/Bradford Books.

Samek-Lodovici, Vieri. 1995. *Constraints on Subjects: An Optimality Theoretic Analysis.* Doctoral Dissertation, Rutgers University.

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

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at the Rutgers Optimality Workshop–1, Rutgers University, October, 1993.

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

Stampe, David. 1979. *A dissertation on natural phonology*. New York: Garland.

Tesar, Bruce B. In press. Error-driven learning in Optimality Theory via the efficient computation of optimal forms. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, eds., *Is the Best Good Enough? Proceedings of the Workshop on Optimality in Syntax*. MIT Press and MIT Working Papers in Linguistics.

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

Tesar, Bruce B., and Paul Smolensky. 1996. Learnability in Optimality Theory. Technical Report, Cognitive Science Department, Johns Hopkins University, Baltimore, Md., and Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, N.J.

Tesar, Bruce B., and Paul Smolensky. Forthcoming. Learnability in Optimality Theory. *Linguistic Inquiry*.

Wexler, Kenneth, and M. Rita Manzini. 1987. Parameters and learnability in binding theory. In Thomas Roeper and Edwin Williams, eds., *Parameter setting*. Dordrecht: Reidel.

Yip, Moira. 1993. Cantonese loan word phonology and Optimality Theory. *Journal of East Asian Linguistics* 2, 262–291.