To appear in *Linguistic Inquiry* 27 (1996)

On the Comprehension/Production Dilemma in Child Language

Paul Smolensky

Johns Hopkins University

The dilemma. That children's linguistic ability in production lags dramatically behind their ability in comprehension poses a long-standing conceptual dilemma for generative studies of language acquisition. Do children's productions reflect their competence in basically the same way as is assumed for adults, or is there a dramatically greater competence/performance gap for children?

The latter hypothesis, invoking severe performance difficulties to account for the impoverishment of production relative to comprehension, has several problems. Gross formulations of the hypothesis, essentially claiming that children don't produce, say, a particular segment because their motor control hasn't yet mastered it, can run afoul of the fact that children who systematically avoid a given structure in their linguistic productions can often easily imitate it (e.g., Menn and Matthei 1992:220). More problematic still for this hypothesis are children like those studied by Smith (1973:149): they produce, for instance, [pʌdəl] and [θɪk], but for *puzzle* and *sick*—*puddle* and *thick* are produced [pʌgəl] and [fɪk].[1] Even subtle formulations of the performance-difficulty hypothesis would seem to entail that generative grammar has little to say about production—in particular, no means of explaining the broad generalization that the additional restrictions manifest in child output align remarkably well with the cross-linguistically observed restrictions on adult outputs: the same configurations which are marked in the sense of disfavored in adult languages tend also to be avoided in child language (Jakobson 1941/1968, Stampe 1979). Where constraints defining linguistic markedness are shared across adult and child language production, and where child productions reveal a grammatical character formally parallel to adult grammars, it would be attractive to have a viable hypothesis according to which grammar has a central role to play in explaining child production.

My topic, therefore, is the other horn of the dilemma. The alternative hypothesis, that the additional limitations manifest in child output are to be explained by a grammar, leads immediately to the extremely unattractive (perhaps, indeed, incoherent) conclusion that the child must have two grammars. It is obvious that the same grammar could not simultaneously yield impoverished productions and relatively rich comprehension.

My purpose here is to prove that this obvious conclusion is incorrect. To provide a minimalist demonstration of this conceptual point, I will present a grammar which simultaneously displays two properties. On the one hand, the grammar leads every word to be produced as the structurally optimal form, perhaps *ba*. On the other hand, this same grammar allows correct comprehension of an unbounded class of words rich in phonemic distinctions and prosodic structure.[2]

This grammar is an Optimality Theoretic one (Prince and Smolensky 1991, 1993). The result presented here contributes to the Optimality Theoretic acquisition literature a general explanation for an assumption central to much of that literature: even children with extremely limited phonological production have underlying forms which relatively closely approximate the adult form (Demuth, in press, Pater and Paradis, in press, Bernhardt and Stemberger 1995, Gnanadesikan 1995, Levelt 1995; see also Smith, 1973). For we shall see how one and the same child grammar can permit the acquisition of a rich set of underlying forms which can be effectively used during comprehension, even though during production most of the underlying distinctions are neutralized to the unmarked structure.

The primary example grammar I will present is a phonological one; the conclusion, however, concerns OT grammars in general, and I include a brief syntactic example as well.

The proposal. The resolution I propose to the comprehension/production dilemma is this: there is only one child grammar, and it is an OT grammar. An OT grammar is to be understood now as a means of evaluating the relative Harmony (unmarkedness) of structural descriptions via a language-particular ranking of universal constraints. These structural descriptions contain two important substructures: the *input* and the *overt form* (in phonology, the underlying and surface forms, respectively). This OT grammar can be deployed in multiple ways. In "production," what is fixed is an input; what competes are structures that share this particular underlying form. The overt expression of a given input is the overt form contained within the maximum-Harmony or optimal structural description of that input. In "comprehension," on the other hand, what is fixed is a surface form, and what competes are structures that share this given overt form. "Comprehension" of a given surface form is determined by the maximum-Harmony structural description containing that surface form. Only one grammar is used in the two processes: one constraint ranking, one means of evaluating the relative Harmony of structures. 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.

As we now see, changing the competition in this way can have dramatic effects on what wins. In "comprehension," it is possible for adult surface forms to be analyzed faithfully: no distinctions are lost. In "production," it is possible for the same grammar to lead to massive loss of distinctions.[3]

In OT terms, this situation at first seems paradoxical: it would appear that faithful "comprehension" requires that FAITHFULNESS constraints be high-ranked, while massive neutralization during "production" requires that FAITHFULNESS constraints be low-ranked. Once the differences in competition are taken into account, however, we see that this apparent paradox dissolves, resolved by prototypical OT constraint-ranking effects.

It should be clear that by "comprehension" and "production" I do not refer to performance, but rather to formal functions of a purely grammatical character, functions that pair underlying and surface forms. The mapping for "production" has been the focus of OT work to date: given an underlying form, which structural description is it paired with, and what is its surface form? The mapping for "comprehension" takes a given surface form, and specifies what structural description it is paired with, and what underlying form is parsed by that structural description. With the understanding that "comprehension" and "production" will refer to entirely competence-theoretic functions, and for want of equally transparent alternative terminology, I henceforth retain the terms but drop the shudder quotes.

The demonstration. The grammar we now investigate is none other than the initial state as loosely proposed in the OT acquisition literature: a hierarchy in which FAITHFULNESS constraints are out-ranked by markedness-defining structural constraints (Demuth, in press, Bernhardt and Stemberger 1995, Gnanadesikan 1995, Levelt 1995; concerning explanation of this initial state from general OT principles, see Smolensky 1996). In the phonological case, the latter constraints include the likes of NOCODA, ‘syllables do not have codas’ (Prince and Smolensky 1991), which determines that closed syllables are marked relative to open ones; *DORS, ‘segments do not have feature [dorsal]’, which determines that dorsal segments (like [k]) are marked (relative to coronal segments like [t], which violate *COR, universally lower-ranked than *DORS; Prince and Smolensky 1993, Smolensky 1993), etc. The details do not matter here: what is crucial is that there is some structure, say, [$_\sigma$ta], which is unmarked relative to other structures (*e.g.*, [$_\sigma$kæt]). With respect to the structural constraints, which I will encapsulate as STRUC-H for *Structural Harmony*, [$_\sigma$ta] is more harmonic than [$_\sigma$kæt].

Potentially conflicting with STRUC-H are the faithfulness constraints encapsulated as FAITH(FULNESS): these demand faithful parsing of the input, and unless the input happens to be /ta/, the demands of FAITH and those of STRUC-H conflict. In the initial grammar, STRUC-H dominates FAITH, so regardless of the input, only the most unmarked structures are produced: [$_\sigma$ta]. This is shown in the tableau under (1); “*” marks constraint violation, “*!” a fatal violation.

(1) Production: /kæt/→ ?

Candidates				Grammar	
	Input	Surface Form	Structure	STRUC-H: NOCODA, *DORS, …	FAITH: PARSE, FILL, …
☞ *a.*	/kæt/	[ta]	σ ⟨k æ⟩ t $\boxed{a}$ / k æ t / [t a]		*
b.	/kæt/	[kæt]	σ k æ t / k æ t / [k æ t]	*!	

These two candidates are alternative parses of the input /kæt/: *b* is a faithful parse, pronounced [kæt], while *a* is pronounced [ta]. In candidate *a*, the first two input segments are unparsed and hence unpronounced; and the segment [a] is inserted in the output in violation of FILL, the faithfulness constraint demanding that syllable positions be filled with underlying material.[4] Parse *a* is favored by STRUC-H (we assume), and since in this grammar STRUC-H dominates FAITH, this is decisive: *a* is optimal (indicated by "☞"), so underlying /kæt/ surfaces as [ta].

What may be less obvious is that this same grammar can be used for comprehension, and that the result is quite different (see tableau (2)).

(2) Comprehension: [kæt]→?

Candidates			Grammar	
Input	Surface Form	Structure	STRUC-H: NOCODA, *DORS, …	FAITH: PARSE, FILL, …
⇛ *b.* /kæt/	[kæt]	σ k æ t / k æ t / [k æ t]	*	
c. /skæti/	[kæt]	σ ⟨s⟩ k æ t ⟨i⟩ / s k æ t i / [k æ t]	*	¡*

A child hearing [kæt], I propose, analyzes this surface form using the grammar in exactly the analogous way to a child producing /kæt/. What is *given* is the surface form, so the competing structures now are all those which are *pronounced* [kæt]: tableau (1)'s winning structure *a* is out of the running altogether. What the faithful parse *b* competes with now are the likes of structure *c*, an unfaithful parse of /skæti/ that is pronounced [kæt]. The violations of STUCT-H incurred by *b* (*e.g.*, *NOCODA), which were fatal in comparison with *a*, are no longer lethal in competition with *c*, which shares the same structural violations: all structures pronounced [kæt] violate NOCODA, *DORS, etc. By standard OT principles, the decision between *b* and *c* is thus passed to low-ranked FAITH, which decisively rules in favor of *b* (here, the fatal violation is marked "¡*" and the winner "⇛", in anticipation of (3)). What we have is McCarthy and Prince’s (1994) *emergence of the unmarked*, with a twist: here, faithfulness constraints have a chance to express themselves when higher-ranked structural constraints do not contravene, the reverse of the pattern extensively documented by McCarthy and Prince.

The contrast between production and comprehension is shown in tableau (3), which combines tableaux (1, 2).

(3) Vive la différence

Candidates				Grammar		Functions using Grammar	
	Input	Surface Form	Structure	STRUC-H: NOCODA, *DORS, …	FAITH: PARSE, FILL, …	Production: /kæt/→?	Comprehension: [kæt]→?
a.	/kæt/	[ta]	σ ⟨k æ⟩ t $\boxed{a}$ / k æ t / [t a]		*	☞	
b.	/kæt/	[kæt]	σ k æ t / k æ t / [k æ t]	*!			←
c.	/skæti/	[kæt]	σ ⟨s⟩ k æ t ⟨i⟩ / s k æ t i / [k æ t]	*	¡*		

FAITHFULNESS, too low-ranked to be active in production, operates decisively in comprehension, despite its low rank. Hearing [kæt], the child analyzes it faithfully as structure *b*—complete with marked segmental structure and marked syllable structure, entirely absent from her productions. This allows her to recognize the underlying form /kæt/ if it is a familiar lexical item, or to enter it into the lexion, if not. (This is a version of Prince and Smolensky's (1993) principle of lexicon optimization; see also Itô, Mester and Padgett 1995 for discussion.)

The proposed relation between production and comprehension can be summarized slightly more formally as follows. Let the universal set of all possible structural descriptions generated by the OT generator of candidates, *Gen*, be the set *UGen*: this set consists of all the candidate parses *Gen*(/ι/) for all universally possible inputs /ι/. Then:

(4) Production and comprehension functions defined

$$f_{\text{prod}}(/\iota/) = \textit{H-max}\{s \in \textit{UGen} \mid /\iota/ = \textit{Input}(s)\}$$

$$f_{\text{comp}}([\text{o}]) = \textit{H-max}\{s \in \textit{UGen} \mid [\text{o}] = \textit{OvertForm}(s)\}$$

The function associated with production takes an input /ι/ and assigns to it the structural description that has the maximum Harmony among the set of all those structures *s* the input part of which is /ι/ (this set is just *Gen*(/ι/), in fact). The function associated with comprehension is exactly the same, except that it operates on an overt form [o] and considers only those structural descriptions *s* the overt part of which is [o].

An illustration from syntax. To better indicate the generality of the current proposal, I briefly present another illustration, based on the approach to *wh*-questions developed in Legendre et al. 1995 and Legendre, Smolensky and Wilson, in press. Tableau (5) shows a sample grammar fragment loosely modeled on a stage of acquisition described in Thornton 1995, when English-learning children produce *medial wh*-phrases, obligatorily with non-referential *wh*-phrases (5*a*: <u>*who*</u> *do you think* ***who*** *the cat chased*), and optionally with referential (discourse-linked) *wh*-phrases (5*a′*,*b′*: <u>*which mouse*</u> *do you think* (***who***) *the cat chased*).

(5) A stage in the acquisition of *wh*-questions (data from Thornton 1995:140)

Candidates			Grammar				Functions using Grammar	
	PF	Structure	MINLINK$^{[-ref]}$	*e	MINLINK$^{[+ref]}$	FILL	Production: Index → ?	Comprehension: (adult) PF → ?
a.	*who* do you think **who** the cat chased	who$_i$ … **wh$_i$** … t$_i$				*	☞	
b.	*who* do you think the cat chased	who$_i$ … **e$_i$** … t$_i$		*!				⇜
c.		who$_i$ … … … t$_i$	¡*!					
a′.	*which mouse* do you think **who** the cat chased	[which mouse]$_i$ … **wh$_i$** … t$_i$				*	☞	
b′.	*which mouse* do you think the cat chased	[which mouse]$_i$ … **e$_i$** … t$_i$		¡*!				
c′.		[which mouse]$_i$ … … … t$_i$			*		☞	⇜

The role of phonology's input/underlying form is now played by what Legendre, Smolensky and Wilson call an *Index*: a predicate/argument structure containing variables with given logical scopes. Candidates (5*a*–*c*) have the same Index, as do (5*a′*–*c′*); the difference is only that the *wh*-variable in the latter is referential whereas that in the former is not. The competing outputs include cyclic structures (*a*, *b*; *a′*, *b′*) with overt (*a*, *a′*) and empty (*b*, *b′*)

intermediate traces, as well as non-cyclic structures (*c*, *c*′). The constraints are as follows (see Legendre, Smolensky and Wilson, in press, for details): *e, forbidding empty intermediate traces; FILL, forbidding more than one overt member of a chain; and MINLINK, which uses OT optimization to force minimal links by forbidding a chain link to cross a barrier. The MINLINK constraints distinguish chains by referentiality, non-referential chains violating universally higher-ranked MINLINK$^{[-\mathrm{ref}]}$. Encapsulating the subtleties of MINLINK, the tableau simply registers that MINLINK favors cyclic chains: candidates *c*, *c*′ respectively violate MINLINK$^{[-\mathrm{ref}]}$ and MINLINK$^{[+\mathrm{ref}]}$. The indicated grammar gives the child production pattern of Thornton 1995 (the two lowest constraints are separated by a dotted vertical line indicating equal ranking: this yields two optimal candidates, and hence optionality).

In comprehension of the adult pronounced form *who do you think the cat chased*, candidate *a* does not compete, since it has the wrong pronunciation: the winner in this non-referential case is now *b*. Even though the child never *produces* empty intermediate traces, her grammar provides an articulated theory of their distribution in adult forms. (The point here is the existence, not the correctness, of such a theory.) Her parse of the non-referential adult PF contains an empty intermediate trace: she gives this chain a cyclic analysis. In contrast, a referential chain is analyzed as non-cyclic: in comprehension of *which mouse do you think the cat chased*, *a*′ is not a competitor; *c*′ wins, because it lacks an empty intermediate trace.

On the role of comprehension in acquisition. The technique proposed above for the use of OT grammars in comprehension is crucial for acquisition theory in other ways. In Tesar and Smolensky 1996, this technique is dubbed *robust interpretive parsing*, and is a central component of an OT learning algorithm. This algorithm starts with an initial grammar: as above, FAITHFULNESS constraints are dominated by structural constraints. The algorithm proceeds by taking overt phonetic forms as primary data, and performing robust interpretive parsing, as above, to assign this data full structural descriptions. Since these descriptions are based on an incorrect grammar, they are not initially correct, but they are used for the next step just the same. The full structural descriptions assigned to the overt data are then used in the Error-Driven version (Tesar, in press) 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, which the algorithm then uses to repeat the whole process over again, reassigning structural descriptions to the primary data and then reranking constraints accordingly. The cycle is iterated repeatedly.

Crucial to this learning procedure and to the arguments presented above is the property of *robustness* of the proposed comprehension process: even when the data being parsed is ungrammatical (suboptimal) according to a learner's grammar, that grammar can nonetheless be effectively used to parse the data. This is what allows a child whose grammar only produces *ta* to correctly parse an unbounded number of phonetic strings, and what enables the learning algorithm to bootstrap a bad grammar into a better one, parsing data which it cannot generate, and using these parses as targets for revising the grammar.

Robust interpretive parsing is essentially an application of Prince and Smolensky's

(1993) principle of *lexicon optimization*, which states that if a learner must choose between alternative underlying forms which generate the correct phonetic form, the one to choose is that which yields the maximally harmonic structural description for that phonetic form. The difference is that here we apply the principle *even when the grammar is incorrect*—even when *no* underlying form will generate the correct phonetic form, due to inadequacies in the grammar. Robust interpretive parsing simply restricts competition to those with the target phonetic form, and picks the one with greatest Harmony according to the current grammar.

Conclusion

A fundamental claim of Optimality Theory is that a grammar is an evaluator of structural descriptions which combines universal constraints—criteria of markedness—to yield a language-particular formal definition of the relative Harmony—unmarkedness—of structures. Because structural descriptions contain both input and overt forms, optimizing Harmony is a principle which defines two related functions from a single grammar. Given either a fixed input or a fixed overt form, optimization assigns a structural description which respectively serves the abstract function of production or comprehension, now considered as part of competence theory. When a child's grammar deviates from the adult grammar, structures which are never produced using the grammar can nonetheless be correctly analyzed, using the same grammar, in comprehension of the adult language.

If grammars are procedures for the sequential derivation of surface forms from underlying forms, grammatically relating production and comprehension is quite difficult; in addition, child grammars become more complex than adult grammars: the least complex derivations are those with output most faithful to input, and the least marked outputs require the most complex derivations to generate. If grammars are parallel optimization over structural descriptions containing both input and surface forms, however, a grammatical relation between comprehension and production is completely natural; furthermore, an initial grammar with low-ranked FAITHFULNESS gives child language production its highly “unfaithful” character, with no increase in grammatical complexity (see also Burzio 1995, Gnanadesikan 1995). If grammars are sets of parametrized inviolable constraints, it is difficult to see how, with a single grammar, children could display one set of parameter settings in their productions, while correctly processing adult forms requiring different settings. If grammars are hierarchies of ranked violable constraints, on the other hand, we expect that children will sometimes correctly process structures they do not produce, because the differences in competitor sets allow constraint interactions masked in production to emerge as decisive in comprehension.

Notes

I am grateful to Stephen Crain, Amalia Gnanadesikan, Norbert Hornstein, Clara Levelt, and Amy Weinberg for stimulating questions and conversations, and to Laura Benua, Luigi Burzio, Géraldine Legendre, Linda Lombardi, John McCarthy, Joe Pater, Colin Wilson, and Lisa Zsiga for very helpful discussion and comments on the manuscript.

1. Thanks to Joe Pater and Linda Lombardi for pointing me to the examples from Menn and Matheii 1992 and from Smith 1973, respectively.

2. In keeping with the minimalist theme, I do not treat the complexities of deriving underlying forms in the face of surface phonological alternations. The proposal developed here actually handles this additional layer of difficulty as well, if, following independently motivated

proposals in the OT phonology literature, optimization is carried out at the level of the morphological paradigm rather than the individual form: see Tesar and Smolensky 1996.

3. It seems likely that accounting for reorganization of perception and comprehension during acquisition will involve additional principles. But under reasonable assumptions, the current proposal need not immediately lead to the conclusion that comprehension is inherently errorless. For example, the loss of the ability to lexically discriminate using features that are non-contrastive in a native language might be consistent with the account presented here, under the additional grammatical assumption that underlying representations are unspecified for predictable material (which follows from one version of lexicon optimization discussed in Prince and Smolensky 1993), and the further extra-grammatical assumption that if a type of linguistic information is not encoded in underlying forms by the end of the critical period for native language acquisition, that type of information cannot later be reliably stored in lexical entries. The proposal of this paper would then entail that a Korean native speaker acquiring English late, when hearing *bear/bail*, will assign the structures [$_\sigma$beːr]/[$_\sigma$beːl], but be unable to reliably distinguish the morphemes in comprehension, because recognition requires matching the assigned structures to stored lexical entries, and the lexical entries for the two items are identical, being unspecified for the r/l distinction. (Thanks to Bob Frank for suggestions along these lines.)

4. What matters about structure *a* is that it is unfaithful and less structurally marked than *b*; only this affects the outcome of competition with *b*, so virtually all the details are irrelevant.

References

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.

Burzio, Luigi. 1995. The rise of Optimality Theory. *Glot International* 1:3–7.

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

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

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

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

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 *Proceedings of the Workshop on Optimality in Syntax—"Is the best good enough?"* Cambridge, Mass: 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 J. and Alan Prince. 1994. The emergence of the unmarked: Optimality in Prosodic Morphology. In *Proceedings of NELS 24*, 333-379. GLSA, University of Massachusetts, Amherst.

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.

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.

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, 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., and Computer Science Department, University of Colorado, Boulder.

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. The initial state and 'richness of the base' in Optimality Theory. Technical Report, Cognitive Science Department, Johns Hopkins University.

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 *Proceedings of the Workshop on Optimality in Syntax—"Is the best good enough?"* Cambridge, Mass: 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.

Thornton, Rosalind. 1995. Referentiality and *wh*-movement in child English: Juvenile *d-link*uency. *Language Acquisition* 4:139–175.