Plenary address: *Proceedings of the 26th Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press (2002).

A Reconsideration of Children's Phonological Representations

Daniel A. Dinnsen
Indiana University

There can be no question that phonological theory has undergone dramatic changes in recent years. For almost forty years, research on fully developed and developing phonologies has been guided by approaches that have been essentially rule-based and derivational. This would include standard generative phonology (e.g., Chomsky & Halle 1968) and natural phonology (e.g., Donegan & Stampe 1979). More recently, however, there has been a shift away from rules and serial derivations toward approaches that are more constraint-based as in optimality theory (e.g., McCarthy & Prince 1995; Prince & Smolensky 1993). The central question that must be asked is whether these approaches are different in ways that should matter to research on acquisition. That is, do any new insights emerge under this newer framework? I want to deal with this question by reconsidering a fundamental issue in acquisition that has occupied us for many years, namely the nature of children's internalized mental representations of words or their underlying representations. The many pronunciation errors that children make in the course of acquiring the target sound system have led us to question whether their underlying representations are more or less like ours. The most widely held assumption has been that children's underlying representations are target-appropriate (or correct), even if their pronunciation of those words is not. There are certainly many good reasons for adopting this assumption. There are also a number of other phenomena and theoretical considerations which have equally well led to the opposite assumption, namely that some children's underlying representations are not target-appropriate. That is, some children's underlying representations would seem to have been internalized incorrectly relative to the target language. Underlying representations could be incorrect in several ways: The set of phonemes available to constitute words might be more restricted. The

---

* I am especially grateful to Judith Gierut, Kathleen O'Connor, Laura McGarrity, Michele Morrisette and Karen Baertsch for their many valuable contributions to all aspects of this work. I would also like to thank Katherine Demuth, Heather Goad and Joe Pater for their input during the conference. This work was supported in part by a grant to Indiana University from the National Institutes of Health DC01694. Contact Information: Department of Linguistics, Indiana University, Memorial Hall 322, 1021 East 3rd Street, Bloomington, IN 47405, E-mail: dinnsen@indiana.edu

substantive representation of phonemes might be more coarsely coded, impoverished, or underspecified. The sequential combinations of phonemes (i.e. phonotactics) might be more limited. Restricting underlying representations in any of these ways has been a perfectly ordinary option that is available for the characterization of a fully developed or developing sound system. However, as we will see, restrictions of this sort on (children's) underlying representations are at odds with a fundamental principle of optimality theory. How then is optimality theory to account for the phenomena that have presumably motivated incorrect underlying representations?

Probably the most compelling evidence in support of children's incorrect underlying representations has come from two characteristic classes of error patterns, namely children's overgeneralization errors and those that might be termed 'complementary error patterns'. I will take up both types of error patterns in the remainder of this paper, illustrating each with a representative case study. While our focus will be on children with phonological delays, it should be kept in mind that these same phenomena have been documented for younger, typically developing children. Different theoretical accounts will be formulated for both types of error patterns with the intent of identifying the different claims and the new insights. We conclude with a brief summary of what optimality theory reveals about these phenomena and acquisition.

1. Overgeneralization errors

Overgeneralization errors have been documented in typical and atypical development and have variously been referred to as recidivism, regressive overgeneralization, hypercorrection, regression, trade-off, or overextension (e.g., Bernhardt & Stemberger 1998; Ingram 1989; Leonard & Brown 1984; Macken 1980; Macken & Ferguson 1983; Menn 1983; Smith 1973). What happens is that a sound that had been produced correctly at one point in time is subsequently replaced by a newly acquired sound, resulting in a new overgeneralization error.

The case study to be considered here was drawn from the Developmental Phonology Archives at Indiana University. The Archives include cross-sectional and longitudinal data on the phonological development of approximately 200 young children with functional (non-organic) speech disorders. Child 78 (age 4;2) was selected to illustrate what we take to be a representative case of overgeneralization. This child produced many sounds in error, scoring at the 5th percentile relative to age-matched peers on the Goldman-Fristoe Test of Articulation (Goldman & Fristoe 1986). She scored within normal limits on all other tests. An extensive speech sample was elicited and revealed that the interdental fricative /θ/ was produced correctly but was also the substitute for all other fricatives.

The data in (1) and (2) from two different points in time are most relevant to the phenomenon of overgeneralization. More specifically, at the first point in time (age 4;2), [s] did not occur in the child's inventory and was replaced by [θ], as shown in (1a). Target [θ] was, however, produced correctly, as shown in (1b).

(1) Stage 1 for Child 78 (age 4;2)

a. Target /s/ replaced by [θ]

[θoʊp]	'soap'	[maʊθ]	'mouse'
[θoʊ]	'sew'	[trihaʊθ]	'treehouse'
[θup]	'soup'	[ʎɛθ]	'yes'

b. Target /θ/ realized as [θ]

[θʌm]	'thumb'	[bæθ]	'bath'
[θʌndʊ:]	'thunder'	[tiθ]	'teeth'

At a second point in time (3 months later), the situation was just the reverse. That is, target [s] came to be produced correctly, as shown in (2a), but [θ] was lost from this child's speech, being replaced by [s], as shown in (2b). The introduction of [s] into the phonology and its correct realization overgeneralized to affect [θ], which had previously been produced without error.

(2) Stage 2 for Child 78 (age 4;5)

a. Target /s/ realized as [s]

[soʊp]	'soap'	[maʊs]	'mouse'
[soʊ]	'sew'	[twihaʊs]	'treehouse'
[sup]	'soup'	[jɛs]	'yes'

b. Target /θ/ realized as [s]

[sʌm]	'thumb'	[bæs]	'bath'
[sʌndʊ]	'thunder'	[tis]	'teeth'

We now turn to a consideration of different theoretical accounts of these phenomena. We begin with the more traditional rule-based account followed by a consideration of the same facts within optimality theory.

Within the more traditional models of phonology, there are in principal several different properties of grammar that are available to account for overgeneralization errors. The most widely held view is that overgeneralization errors result from the combination of rule loss and incorrectly internalized underlying representations (e.g., Macken 1980). Adopting this approach to account for the facts of Child 78, it might be argued that during the two stages all of her fricatives were restricted underlyingly to one undifferentiated fricative category, i.e., /s/. Target /s/ would thus have been represented more or less correctly, but the other fricatives would have been internalized incorrectly as /s/. The realization of all fricatives as [θ] at the early stage would be achieved by a rule or natural process as in (3).

(3) Dentalization rule for Stage 1

[+consonantal, +continuant] → [coronal, -strident]
(All fricatives are realized as [θ].)

With this rule, target /θ/ would be realized correctly, albeit from an incorrectly internalized underlying representation. The sample derivations in (4) illustrate the assumptions and results in Stage 1 for three words beginning with different target fricatives. It should be noted that the underlying representation for all of these words would begin with the fricative /s/. The Dentalization rule in (3) would then convert that fricative to the actually occurring forms with [θ].

(4) Sample derivations for Stage 1

UR	/soup/ 'soap'	/sʌm/ 'thumb'	/sæt/ 'fat'
Rule (3)	θoup	θʌm	θæt
PR	[θoup]	[θʌm]	[θæt]

The subsequent target appropriate realization of /s/ along with its overgeneralization at Stage 2 would follow simply from the child's loss or suppression of her Dentalization rule in (3). The absence of the rule from the grammar would permit all fricatives to be realized as they were internalized underlyingly. The overgeneralization errors thus serve to reveal the child's previously incorrect underlying representations for various words. Under such an account, the underlying representations would remain unchanged throughout the two stages. Those underlying representations would have been correct for some fricatives (i.e., for /s/), but incorrect for others (for /f/ and /θ/) at both points in time. Such an account entails restrictions on the child's underlying representations and attributes the change over time to rule loss. The sample derivations in (5) illustrate these points for Stage 2 using the same three words.

(5) Sample derivations for Stage 2

UR	/soup/ 'soap'	/sʌm/ 'thumb'	/sæt/ 'fat'
Rule (3) lost	----	----	----
PR	[soup]	[sʌm]	[sæt]

The above account suggests that at the point overgeneralization errors occur, the learner will still need to restructure the underlying representation of many words to bring her system into conformity with target English. Independent of the merits of this account, it does at least serve to illustrate one derivational mechanism that is available to yield overgeneralization. In fact, any error pattern could be attributed to incorrect underlying representations and a rule, which when lost could yield overgeneralization errors. Given this possibility, we might expect overgeneralization phenomena to be unconstrained, occurring freely in association with any or all error patterns in the course of development. It remains to be determined whether this expectation is empirically warranted. In what follows, we will see that optimality theory makes different predictions about the occurrence and nonoccurrence of overgeneralization errors.

Before embarking on our optimality theoretic account of these facts, we provide a brief sketch of the essentials of this newer framework. For a tutorial introduction to optimality theory with special reference to acquisition, see Barlow & Gierut (1999) and Kager (1999). Optimality theory differs from derivational theories in several important respects. Some of the central hypotheses are that there are no rules, no serial derivations, no intermediate levels of representation, and no language-specific restrictions on the set of available input representations (underlying representations). Instead, for any given input representation, a ranked set of universal constraints evaluates in parallel a potentially infinite set of output candidates and selects one as optimal. The optimal phonetic form is the one that best satisfies the constraint hierarchy. The constraint hierarchy is a language-specific ranking of the constraints. Languages are presumed to differ solely by the ranking of constraints.

Constraints are of two fundamental and often antagonistic types, namely markedness constraints and faithfulness constraints. Markedness constraints are formulated exclusively in terms of output properties without regard to the input representation and militate against relatively marked outputs. Faithfulness constraints, on the other hand, demand identity between corresponding elements in input and output representations. Faithfulness constraints militate against output candidates that are different from an input representation. Faithfulness constraints are thus the antithesis of rules in that they disfavor change. Another important difference between constraints and rules is that the constraints are presumed to be universal (whereas rules are language-specific). The constraints are thus the same across languages and are present in all grammars. The conflict between constraints is resolved by constraint rankings. Some constraints will dominate or outrank other constraints.

In terms of acquisition concerns, it is hypothesized that markedness constraints outrank faithfulness constraints in the early stages of acquisition (e.g., Smolensky 1996a, b). The process of acquisition is presumed to proceed by constraint demotion (e.g., Tesar & Smolensky 1998).

Children's overgeneralization errors pose a special challenge for optimality theory. Because there are no rules within this framework, there can be no rule loss, what had been an essential property of derivational accounts of overgeneralization errors. Additionally, while a derivational account might attribute to a child incorrectly internalized underlying representations by limiting those representations to, for example, a single fricative category /s/, optimality theory prohibits restrictions on the set of available input representations. This optimality theoretic principle, known as 'richness of the base', maintains that the set of underlying representations is universal and thus the same for all languages and by extension for all children (e.g., Smolensky 1996a). Any account of these overgeneralization errors in this framework must therefore provide for the possibility that the child's underlying representations could be target appropriate.

We now turn to an optimality theoretic account of Child 78's error patterns. The constraints relevant to this case are given in (6) along with their associated definitions.

(6) Constraints and definitions

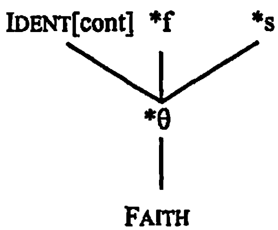
a. Markedness constraints
   - *s: Avoid strident coronal fricatives
   - *θ: Avoid interdental fricatives
   - *f: Avoid labial fricatives

b. Faithfulness constraints
   - IDENT[cont]: Preserve the feature [continuant]
   - FAITH: Preserve all features

The markedness constraints in (6) all belong to a family of constraints disfavoring fricatives generally. Each individual constraint militates against a different class of fricatives and each is independently necessary to account for individual differences in the occurrence and non-occurrence of particular fricatives across children. The individual differences that we are referring to relate to the fact that some children exclude all fricatives from their inventories, others exclude only one class or some combination of those classes, and yet others exclude none. While many fricatives were banned from Child 78's inventory, it is noteworthy that the substitute for all target fricatives was a fricative, specifically an interdental fricative at Stage 1. This is suggestive of a highly ranked faithfulness constraint, IDENT[cont], which demands that the input manner feature [continuant] be preserved in the corresponding output segment. The dominance of this constraint would ensure that a target stop is realized as a stop and a target fricative as a fricative. Child 78's exclusion of labial fricatives and strident coronal fricatives is indicative of the highly ranked markedness constraints *f and *s, respectively. By ranking these two markedness constraints over another general faithfulness constraint, FAITH, which demands that all properties of an input and output be the same, it is claimed that it is more important to avoid labial fricatives and strident coronal fricatives than it is to preserve their place and stridency features. The consequence of these various constraints and rankings would be that labial fricatives and strident coronal fricatives would be excluded from the inventory and replaced by the only remaining class of English fricatives, namely interdental fricatives. It is, however, also well known that many children exclude interdental fricatives from their inventories, suggesting the need for the additional independent markedness constraint, *θ, which disfavors interdental fricatives. Given that Child 78 produced interdental fricatives target appropriately at Stage 1 and, in fact, preferred interdentals as the substitute for all other fricatives, the first thought would be that *θ is dominated by FAITH. Such a ranking would ensure the target appropriate realization of /θ/ and would assess less serious (non-fatal) violation marks to occurring forms with interdentals. We will see, however, that learnability and continuity considerations require that *θ be ranked above FAITH but below the other markedness constraints during the early stage.

The ranking of constraints needed for Stage 1 is schematized in (7). Solid lines connect those constraints that are crucially ranked with the higher ranked constraints positioned above the lower ranked constraints. Constraints that cannot be ranked relative to one another are given on the same horizontal plain and are not connected by a line.

(7) Constraint ranking for Stage 1



With this ranking of constraints, we can now demonstrate how a particular output candidate is selected as optimal given a specific input (or underlying representation). It is conventional to use a display known as a tableau for this purpose. Our first tableau is given in (8). In all tableaux, the input representation is given in the upper left corner. Competing output candidates are listed down the left side of the tableau. For expository purposes, we will limit the output candidate set to the most likely competitors, namely those that begin with a fricative and differ in their place and stridency features. It is assumed that all other competing output candidates would be eliminated by other highly ranked constraints (e.g., IDENT[cont]) that need not be detailed here and will not be shown in the tableaux. Constraints are listed along the top in accord with their ranking. Crucial rankings are indicated by a solid vertical line between affected constraints. Constraints whose ranking cannot be determined are separated by a dotted vertical line. A candidate's violation of a constraint is indicated by a '*' in the intersecting cell. The elimination of a candidate from the competition is termed a fatal violation and is indicated by a '!' after the violation mark. The winning or optimal candidate is identified by the manual indicator '☞'.

The tableau in (8) considers how different output candidates would be evaluated given this ranking of constraints and an input representation that begins with /s/ as in 'soap'. Candidate (a) with an initial [f] fatally violates the highly ranked markedness constraint *f and is eliminated from the competition. The faithful candidate (c) violates the other highly ranked markedness constraint *s and is also eliminated. Candidate (b) with the interdental is all that remains and is thus selected as optimal in accord with the child's error pattern even though it violates both *θ and FAITH. The lower ranking of those two constraints makes their violations less serious. Preservation of input features is apparently not decisive in this instance. For this reason, candidate (b) would also be selected as optimal even if the input representation began with a labiodental fricative. Consequently, with this ranking of the constraints, the substitution of [θ] for both classes of fricatives can be achieved independent of the particular fricative that might have been internalized by this child. This accords well with richness of the base because no restrictions are being imposed on the set of input representations.

(8) Target /s/ realized as [θ] for Stage 1

/soup/ 'soap'	*f	*s	*θ	FAITH
a. [foup]	*!			*
b. ☞ [θoup]			*	*
c. [soup]		*!		

Legend for tableaux:

☞ = optimal output   * = constraint violation   ⋮ = equal ranking

*! = fatal violation | = crucial ranking

This same ranking of constraints ensures the target appropriate realization of /θ/ as shown in the tableau in (9) for an input such as 'thumb'. Candidates (a) and (c) are again eliminated because each violates one of the undominated markedness constraints. The only remaining candidate (b) complies with FAITH and is selected as optimal even though it violates the markedness constraint *θ.

(9) Target /θ/ realized as [θ] for Stage 1

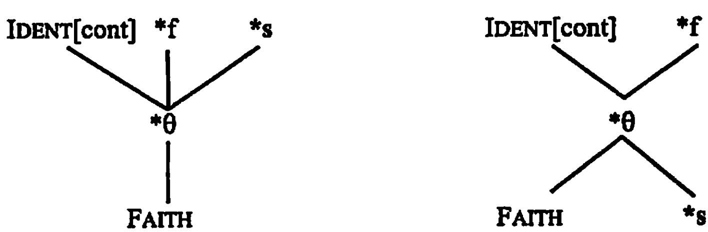
/θʌm/ 'thumb'	*f	*s	*θ	FAITH
a. [fʌm]	*!			*
b. ☞ [θʌm]			*	
c. [sʌm]		*!		*

Ranking all of the markedness constraints above FAITH as we have done here accords well with the assumption that markedness constraints tend to outrank faithfulness constraints in the early stages of acquisition. In addition, by ranking the markedness constraints relative to one another, specifically both *f and *s above *θ, all competing fricative candidates of English, except for the interdental fricatives, are effectively eliminated. Thus, even though *θ is ranked above FAITH (seemingly disallowing interdentals), interdental fricatives are permitted to survive as optimal because of the greater demand to preserve the manner of the input segment in the corresponding output segment (i.e,. as a result of undominated IDENT[cont]). The claim is that interdental fricatives are dispreferred, but they are better than no fricative at all for an input fricative. They are the fricative of last resort in this instance. It turns out that target-appropriate realizations of interdentals have little to do with FAITH or any assumptions about the place or stridency of the input fricative. It appears that target /θ/ was produced correctly for the wrong reasons. That is, [θ] would have been the realization for target /θ/ no matter what fricative the child internalized in the underlying representation. Also, correct realizations of /θ/ resulted from a constraint ranking that does not conform to the target ranking of constraints, namely in adult English FAITH presumably dominates *θ.

Let us now turn to the characterization of Stage 2 and its transition from Stage 1. Given the ranking of constraints for Stage 1 (repeated in (10)), a natural account becomes available for the transition to Stage 2 with its target-appropriate realizations of /s/ and the overgeneralization errors associated with the other fricatives. More specifically, all that has to happen is for *s to be demoted minimally below *θ, resulting in the new ranking for Stage 2 as shown in (10). This reranking is precisely what would follow from the constraint demotion algorithm as put forward by Tesar & Smolensky (1998). That is, upon the child's recognition that [s] can occur, she would have had to demote the constraint responsible for the exclusion of [s], namely *s, just below the highest ranked constraint that her previous winner (b) violated, in this case just below *θ. This has the effect of installing *s in the stratum with FAITH. Everything else remains the same across the two stages.

(10) Constraint demotion for Stage 2

Stage 1: Stage 2:



The consequence is that the ranking of *θ over FAITH would be preserved in the transition from Stage 1 to Stage 2, accounting for the new error pattern where [θ] was dispreferred and replaced by [s]. Also common to both stages is the continued dominance of *f, which accounts for the persistent exclusion of labiodental fricatives from the child's inventory.

The tableau in (11) shows how the ranking in (10) for Stage 2 ensures target appropriate realizations of /s/ in words such as 'soap'. The faithful candidate (c) only violates the lowest ranked markedness constraint *s with all of the competitors being eliminated by their violations of the higher ranked markedness constraints.

(11) Target /s/ realized as [s] for Stage 2

/soup/ 'soap'	*f	*θ	FAITH	*s
a. [foup]	*!		*	
b. [θoup]		*!	*	
c. ☞ [soup]				*

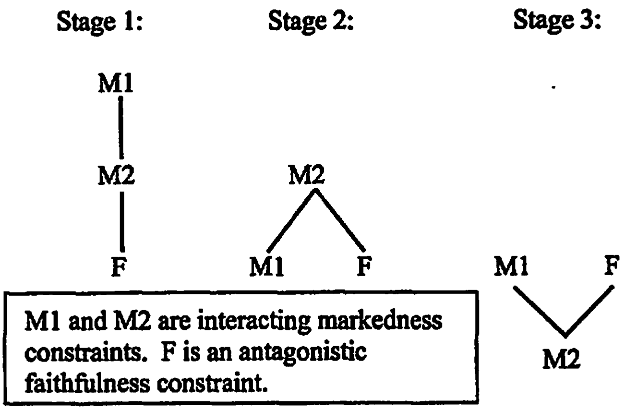
The tableau in (12) illustrates how the constraint ranking in (10) for Stage 2 yields the overgeneralization error where previously correct /θ/ came to be replaced by [s]. Given a target word such as 'thumb', candidates (a) and (b) are both eliminated by their fatal violations of the more highly ranked markedness constraints. Candidate (c) with the substitute [s] is selected as optimal in accord with the child's new error pattern even though that candidate violates the two lower ranked constraints. This also explains in part why the substitute for /f/ would have at the same time changed from [θ] to [s]. That is, the higher ranking of *f and *θ precludes the occurrence of candidates (a) and (b) independent of the particular fricative that might have been internalized by the child.

(12) Target /θ/ realized as [s] for Stage 2

/θʌm/ 'thumb'	*f	*θ	FAITH	*s
a. [fʌm]	*!		*	
b. [θʌm]		*!		
c. ☞ [sʌm]			*	*

The general characteristics of our account can be extended to other cases of overgeneralization. The schematization in (13) sets out in optimality theoretic terms the essentials that give rise to overgeneralization errors and the associated course of development.

(13) Schema for overgeneralization errors and stages of development



According to this schema, overgeneralization errors arise from a characteristic set of grammatical factors and proceed in several crucial developmental steps. The first step in the developmental progression entails an early stage where an error pattern is characterized by two highly ranked and interacting markedness constraints, M1 and M2. These markedness constraints are ranked relative to one another and are both ranked above an antagonistic faithfulness constraint, F.[1] These markedness constraints can be observed to interact by virtue of their conflict with one another in overlapping contexts. More precisely, different rankings of the markedness constraints relative to one another would have different empirical consequences (i.e., result in different error patterns). The second step in the developmental progression entails the reranking of the markedness constraints such that M1 is demoted minimally below M2. However, because M2 would continue to outrank F, its influence would then become evident with the introduction of a new error pattern. This is entirely consistent with Tesar and Smolensky's (1998) constraint demotion algorithm in that the reranking is motivated on the basis of positive evidence alone, namely the child's recognition that the sounds disfavored by M1 actually do occur. The error pattern induced by the dominance of M2 during Stage 2 means that a further change must take place before conformity with the target system can be achieved. Stage 3 is predicted to be the next stage of development approximating the target language. That is, a further round of constraint demotion is needed whereby M2 is demoted below F. The ranking for Stage 3 would allow the sounds disfavored by M1 and M2 to both occur.

A further distinguishing characteristic of this optimality theoretic account is that the substance of a child's underlying representations plays a much smaller role and is less decisive than might have been the case in prior accounts. That is,

1. It must be kept in mind that a higher ranked faithfulness constraint (e.g., IDENT[cont]) preserves some property of the sounds disfavored by M1 and M2.

it is not necessary to assume that the child's underlying representations are correct or incorrect relative to the target system during either of the first two stages of development. Naturally, this should not be taken to the absurd extreme--that nothing about the target representations matters. We saw, for example, in the case of Child 78 that, although target place features did not matter, the manner of target obstruents did as captured by the dominance of IDENT[cont]. The fact that overgeneralization errors involve correct and incorrect productions during the first two stages of development has less to do with the substance of underlying representations and more to do with the nature of the constraints and the constraint hierarchy. In these circumstances, optimality theory views overgeneralization errors as a natural and unavoidable intermediate step in the right direction, but one that may require a series of constraint rerankings to achieve conformity with the target language.

We now turn to the other class of error patterns, namely complementary errors, which have also been claimed to reflect a problem with learners' underlying representations.

2. Complementary error patterns

Complementary error patterns have been documented in both typical and atypical development (e.g., Camarata & Gandour 1984; Dinnsen 1996, 1999; Gierut 1986, 1989; Gierut & Champion 2000; Smith 1973; Williams & Dinnsen 1987). These error patterns are also especially common, and perhaps better understood, in second-language acquisition, where a learner transfers his/her native language allophonic rule into the interlanguage (e.g., Eckman & Iverson 1997; Hardy 1993; Lado 1957; Sheldon & Strange 1982). Transferring this rule results in the complementary distribution of sounds that should be separate phonemes in the target language. Because the rule prevents the target sounds from occurring in the same context, they cannot contrast, and systematic errors are introduced. The complementary distribution of sounds is, of course, a highly common and natural phenomenon in the languages of the world. It thus should not be surprising that this same phenomenon might occur in the developing phonologies of young children, even if it results in an error pattern.

When sounds are observed to occur in complementary distribution in any language system, it is typically assumed that those sounds are simply allophones of a single phoneme and that the different contextual realizations of that phoneme are a predictable property of pronunciation. Consequently, when children exhibit complementary error patterns, it is assumed that only a subset of the target phonemes is available to constitute the underlying representation of words. Stated differently, complementary errors are taken as indicative of fewer phonemes than in the target language.

We have selected the case study of another child with phonological delays to illustrate what we take to be a representative example of a complementary error pattern and its associated course of development. The case of AJ (age 4;11) as reported by Gierut (1986, 1989) (cf. Dinnsen & Chin 1995; Fey 1989) has attracted attention from various theoretical perspectives. This case is important because three stages of development have been identified in the transition from a complementary error pattern to something approaching the target system where the sounds are properly associated with separate phonemes (i.e., a phonemic split).

AJ's presenting phonology included only two fricatives, [f] and [s]. These two fricatives served as substitutes for the other English fricatives and were produced correctly in certain well-defined contexts but incorrectly in others. More specifically, [f] was produced correctly in word-initial position but incorrectly in other contexts; [s] was produced correctly in post-vocalic contexts but incorrectly elsewhere. Independent of target considerations, these two fricatives were observed to occur in complementary distribution with [f] occurring exclusively in word-initial position and [s] occurring elsewhere. These distributional restrictions are exemplified by the data in (14).

(14) Stage 1 for AJ (age 4;11): Complementary distribution

a. [f] limited to word-initial position

[fæt]	'fat'	[fæni]	'funny'
[fu]	'food'	[fes]	'face'

b. [s] limited to post-vocalic position

[ɔs]	'off'	[maʊs]	'mouse'
[gusi]	'goofy'	[aʊs]	'house'

AJ's two fricatives clearly resembled the two separate phonemes /f/ and /s/ of English, although they did not appear to contrast in his speech. His distributional restrictions on the occurrence of these sounds precluded a phonemic contrast, suggesting that they were allophones of a single phoneme. The claim would be that AJ's phonemic inventory at Stage 1 was more restrictive than in English, being limited to one fricative phoneme. The choice of the fricative to serve as that one phoneme would seem to be relatively arbitrary given that either could be derived from the other by a simple rule. We will assume that AJ represented all fricatives by the labial phoneme /f/. An allophonic rule similar to that in (15) would account for the complementary distribution of the fricatives by converting the one fricative phoneme into [s] in post-vocalic contexts, leaving it unchanged in all other contexts.

(15) Allophonic rule for Stage 1

[-sonorant, +continuant] → [coronal] / V __
(Fricatives are realized as [s] post-vocalically and as [f] elsewhere.)

The derivations in (16) illustrate the relevant assumptions about the underlying representation of fricatives and the operation of the rule for Stage 1. Special attention is drawn to the assumption that all target fricatives were represented underlyingly as /f/. This entails a correct assumption for some words (e.g., 'fat' and 'off') but an incorrect assumption for certain other words (e.g., 'mouse'). Given the contextual restrictions on the rule, word-initial target /f/'s would not be affected and would be realized correctly (e.g., 'fat'). This is in contrast to target /f/'s in post-vocalic contexts, which would be represented correctly but would also undergo the rule, resulting in an error (e.g., 'off' realized as [ɔs]). A further point of interest is the correct result that obtains for post-vocalic target /s/, which derives from the operation of the rule on an incorrect underlying representation (e.g., 'mouse' from a final /f/).

(16) Sample derivations for Stage 1

UR	/fæt/ 'fat'	/ɔf/ 'off'	/maʊf/ 'mouse'
Rule (15)	----	ɔs	maʊs
PR	[fæt]	[ɔs]	[maʊs]

Given the anomalies of AJ's phonology, a clinical intervention plan was devised and implemented with the intent of splitting the two allophones [f] and [s] into separate phonemes. Accordingly, these two sounds were opposed to one another in minimal pairs, specifically in post-vocalic position. This reportedly resulted in only moderate success. That is, while the contrast between /f/ and /s/ began to emerge in the treated context, it certainly did not extend to all relevant words. Additionally, as shown in (17), the new post-vocalic [f]'s freely varied with [s]. For example, target words such as 'off' were realized as [ɔf] and [ɔs]. It is noteworthy that target /s/'s in that same context did not exhibit variation (e.g., 'mouse' was produced consistently with a final [s]).

(17) Stage 2 for AJ (age 5;0)

[ɔf]	~	[ɔs]	'off'
[gufi]	~	[gusi]	'goofy'
[gɔfɪn]	~	[gɔsɪŋ]	'golfing'
[naɪfi]	~	[naɪsi]	'knifey'

These results were taken as characteristic of a second, intermediate stage of development in the transition from a complementary error pattern to a full phonemic split. The facts of this stage are suggestive of several changes in AJ's phonology. On the one hand, the number of fricative phonemes would have increased from one to two given that /f/ and /s/ began to contrast. The other change in AJ's phonology was that the newly introduced contrast between post-vocalic /f/ and /s/ was optionally being neutralized. This means that the earlier obligatory allophonic rule (15) changed to a neutralization rule in Stage 2 and became optional. The sample derivations in (18) illustrate the change in the underlying representation of target /s/'s but not /f/'s post-vocalically along with the optional merger of /f/ and /s/ in that context. In the case of target words such as 'off', the optional rule may or may not apply. In the case of target words such as 'mouse', the restructured and correct underlying representation with an /s/ renders the rule vacuous.

(18) Sample derivations for Stage 2

UR	/ɔf/ 'off'	/ɔf/ 'off'	/maus/ 'mouse'
Rule (15)	ɔs	---	-----
PR	[ɔs]	[ɔf]	[maus]

Given this variation and the persistence of the original error pattern in so many words and especially in initial position, an additional round of minimal pair treatment was deemed necessary, focusing on the contrast between /f/ and /s/ word-initially. This round of treatment resulted in a third stage of development with a number of changes--not all of which seem positive. While the contrast did extend to more words, some of the words that had been produced correctly at Stage 2 reverted back to their original error pattern. That is, some post-vocalic target /f/'s returned to the original error pattern, being realized as [s] (e.g., 'off' was realized as [ɔs]). It is unclear whether this would have been the result of a change in the underlying representation of those words or in the operation of the rule.

There are at least three unexplained peculiarities about this case that must be accounted for by any adequate theory of phonology. First, one striking fact that distinguishes this case from that of Child 78 is that, as the new sounds emerged in their new contexts, no overgeneralization errors occurred. That is, [f] never came to replace target /s/ in post-vocalic contexts, nor did [s] ever come to replace target /f/ in word-initial position. Another peculiarity is the asymmetry that occurred with respect to those sounds that showed free variation during Stage 2. That is, why did target /f/ in post-vocalic contexts alternate with [s], but target /s/ did not alternate? Finally, why did some of the correct productions from Stage 2 revert back to the original error pattern at Stage 3? We will see in what follows that optimality theory offers answers to these questions.

Our optimality theoretic account of this error pattern and its associated stages of development leading to a phonemic split will be shown to follow from a series of constraint rerankings with no restrictions on the child's underlying representations. The constraints relevant to this case are given in (19). These constraints expand a bit on those relevant to the other case study. The most significant difference relates to the explosion or unpacking of the markedness constraints to account for the context effects. That is, while coronal and labial fricatives are each disfavored by general markedness constraints, *s and *f, respectively, each of these constraints has a more specific, contextually-conditioned counterpart that assesses more serious violations for the occurrence of these sounds in particular contexts. Thus, while coronal fricatives might generally be disfavored (*s), the claim of the positionally-restricted constraint (*#s) is that coronal fricatives are especially disfavored in word-initial contexts. Similarly, while labial fricatives might in general be dispreferred (*f), the claim of the other positionally-restricted constraint (*Vf) is that labial fricatives should especially be avoided in post-vocalic contexts. The motivation or rationale for contextual restrictions of this sort is likely related to considerations of prominence or salience. For example, the post-vocalic context of *Vf essentially defines foot-medial and foot-final positions, which are acknowledged to be weak and thus not good contexts for a low intensity fricative such as [f]. A similar restriction holds for fully developed languages (e.g., Spanish) where the only fricative phoneme allowed in foot-final position is the high intensity fricative /s/. It is less clear why [s] should be dispreferred in word-initial position, unless possibly the enhanced visual salience of [f] plays some role especially in early phonological development. A further assumption is that each of the contextually-conditioned constraints universally outranks its context-free counterpart. It should be underscored that the contextually-conditioned constraints militate against different classes of fricatives in non-overlapping environments and thus could not interact with one another no matter how they might be ranked. Any of these markedness constraints could, however, interact with the general faithfulness constraint, FAITH, which demands identity between corresponding segments, especially in terms of place features.

(19) Constraints

a. Markedness constraints

*#s:	Avoid coronal fricatives word-initially
*s:	Avoid coronal fricatives (elsewhere)
*Vf:	Avoid labial fricatives post-vocalically
*f:	Avoid labial fricatives (elsewhere)

b. Faithfulness constraint

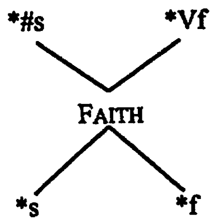
FAITH: Preserve all features

The complementary distribution facts of Stage 1 can be accounted for by ranking the two contextually-conditioned markedness constraints above FAITH as shown in (20).[2] Inasmuch as those two markedness constraints are defined on

2. The same result would obtain for Stage 1 if all of these markedness constraints were alternatively ranked above FAITH. We will see, however, that the characterization of the subsequent stages of development depends on the lower ranking of the context-free markedness constraints as shown in (20).

nonoverlapping contexts, no different result would follow from ranking either constraint over the other; the two are thus unranked relative to one another.

(20) Ranking for Stage 1 with complementary distribution



With the ranking in (20), it is not necessary to make any particular assumption about which fricative phoneme or phonemes might underlie AJ's error pattern. The next three tableaux show how the complementary distribution of these fricatives follows from different input representations. In each case, one of the two top-ranked constraints eliminates one of the two most likely competitors.

(21) Word-initial target /f/ realized as [f] for Stage 1

/fæt/ 'fat'	*#s	*Vf	FAITH	*s	*f
a. ☞ [fæt]					*
b. [sæt]	*!		*	*	

(22) Post-vocalic target /f/ realized as [s] for Stage 1

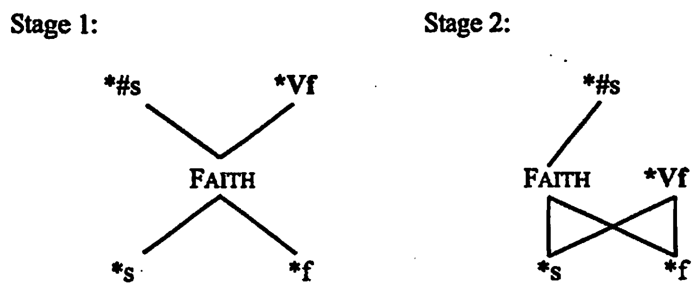
/ɔf/ 'off'	*#s	*Vf	FAITH	*s	*f
a. [ɔf]		*!			*
b. ☞ [ɔs]			*	*	

(23) Post-vocalic target /s/ realized as [s] for Stage 1

/maʊs/ 'mouse'	*#s	*Vf	FAITH	*s	*f
a. ☞ [maʊs]				*	
b. [maʊf]		*!	*		*

To account for the facts of Stage 2, all that is available is the reranking of the constraints. The emergence of the phonemic contrast and its optional neutralization in post-vocalic contexts in Stage 2 can be seen to follow from the minimal demotion of a single contextually-conditioned markedness constraint, namely *Vf, as shown in (24). The other ranking relationships from Stage 1 would remain unchanged.

(24) Constraint demotion for Stage 2



The demotion of *Vf into the stratum with FAITH results in a tie between candidates corresponding to a post-vocalic target /f/, as shown in the tableau in (25). The faithful candidate (a) violates the contextually conditioned markedness constraint, while candidate (b) violates FAITH. The choice between the two candidates must then be passed down to the context-free markedness constraints, but again each candidate violates one of the equally ranked constraints, resulting in the observed free variation for target /f/.

(25) Optional neutralization for Stage 2

/ɔf/ 'off'	*#s	FAITH	*Vf	*s	*f
a. ☞ [ɔf]			*		*
b. ☞ [ɔs]		*		*	

The sustained absence of variation in words with post-vocalic target /s/ is achieved by the same ranking in (24) for Stage 2 and is illustrated in the tableau in (26). The candidate in (b) is decisively eliminated by its violation of FAITH (and/or *Vf); the faithful candidate (a) violates neither of the constraints in that stratum and thus survives as optimal.

(26) Non-alternating post-vocalic /s/ for Stage 2

/maʊs/ 'mouse'	*#s	FAITH	*Vf	*s	*f
a. ☞ [maʊs]				*	
b. [maʊf]		*!	*		*

The extension of the phonemic contrast to word-initial position in Stage 3 is achieved by the demotion of the other contextually-conditioned markedness constraint, namely *#s, below FAITH, being installed into the same stratum with the context-free markedness constraint *f, as shown in (27). The fixed universal ranking of a context-sensitive constraint over its context free counterpart forces the further demotion of *s.

(27) Constraint demotion for Stage 3

Stage 2: Stage 3:

The dominance of FAITH in this instance permits the phonemic contrast between /f/ and /s/ to be realized target appropriately in word-initial position. However, the demotion of *#s has an interesting consequence for the contrast in post-vocalic contexts. Note that the constraint hierarchy for Stage 3 retains the equal ranking of FAITH and *Vf from Stage 2. As a result, target /f/'s in post-vocalic contexts will appear to have regressed back to the original error pattern being replaced by [s]. The tableau in (28) for the target word 'off' illustrates this effect.

(28) Regression of post-vocalic /f/ in Stage 3

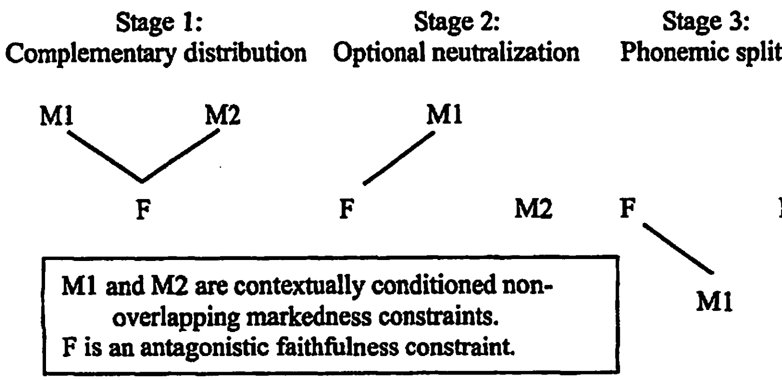
/ɔf/ 'off'	FAITH	*Vf	*#s	*f	*s
a. [ɔf]		*		*!	
b. ☞ [ɔs]	*				*

The equal ranking of FAITH and *Vf yields a tie between candidates (a) and (b). The choice is then passed down to the next lower ranked constraints. Because the two context-free markedness constraints are no longer in the same stratum, the faithful candidate (a) incurs a fatal violation of *f and is eliminated in favor of candidate (b). This apparent regression was a rather surprising result in the traditional derivational account, but it is the expected result within optimality theory due to the minimal and incremental rerankings associated with constraint demotion. While we have no evidence about AJ's progress beyond Stage 3, it appears likely that he would achieve conformity with the target system upon the simple demotion of *Vf below FAITH. That demotion might require yet another round of treatment that focused once again on the /f/ - /s/ contrast in post-vocalic contexts.

In contrast to a traditional derivational account and consistent with richness of the base, our optimality theoretic account of AJ's complementary error pattern and his course of development in effecting a phonemic split was accomplished without imposing any restrictions on his underlying representations. A number of other previously unexplained peculiarities associated with the different stages of development also fell out as a natural consequence of constraint demotion. In particular, we were able to explain why during Stage 2 it was the post-vocalic target /f/'s, but not /s/'s, that exhibited free variation. It also followed that it would be that same class of words that appeared to have regressed during Stage 3. Finally, our account provides for the absence of overgeneralization errors in the course of effecting a phonemic split. That is, there is no one ranking of the constraints that would predict word-initial target /f/ to be realized as [s] while at the same time predicting post-vocalic target /s/ to be realized as [f].

The general characteristics of our account should apply equally to other complementary error patterns as a phonemic split evolves and can be schematized as in (29).

(29) Schema for complementary error patterns and phonemic splits



According to this schema, complementary error patterns are characterized by two markedness constraints that disfavor different sounds in non-overlapping contexts.[3] Different rankings of these markedness constraints relative to one

3. It is acknowledged that a rather different schema for allophonic phenomena has generally been assumed within optimality theory, at least for fully developed languages (e.g., Kager 1999). That schema entails a context-sensitive markedness constraint that is ranked above a context-free markedness constraint, where the two constraints refer to different sounds in overlapping contexts and can interact. Such a characterization would be insufficient as an

another could have no different empirical consequence and thus could not directly interact. This is the essential property that distinguishes complementary error patterns from other error patterns that are vulnerable to overgeneralization. That is, overgeneralization errors come about from two or more markedness constraints that disfavor different sounds in overlapping contexts. Consequently, different rankings of those constraints relative to one another will yield different, seemingly contradictory results. These same optimality theoretic considerations have been applied with equal success to the characterization of overgeneralization errors and complementary errors in normal development (e.g., Dinnsen, O'Connor & Gierut 2001).

3. Conclusion

Let me return to our original question: Does optimality theory tell us anything new about acquisition? I submit that the answer is 'yes', especially as revealed through a reconsideration of children's phonological representations. At the very least, one new insight that emerges from optimality theory is its characterization of children's error patterns. More specifically, the characterization of overgeneralization errors and complementary errors has identified important commonalities and differences. In terms of commonalities, while traditional derivational theories have argued that overgeneralization errors and complementary error patterns require restrictions on children's underlying representations, we have seen, consistent with richness of the base, that optimality theory can account for these facts without imposing any such restrictions. A second unifying property of these error patterns is their reliance on multiple independent markedness constraints that outrank an antagonistic faithfulness constraint. That is, overgeneralization errors and complementary errors are really seen to be the product of multiple error patterns. Related to this point is the further observation that multiple constraint rerankings or demotions are necessary to eradicate overgeneralization errors and complementary errors. This characterization represents a significant advance because we now have something of an explanation for why these two types of error patterns are so much more resistant to change, especially compared with other types of error patterns. That is, many other error patterns can be characterized as the product of a single markedness constraint that outranks a faithfulness constraint. The elimination of these other error patterns would simply require a single constraint reranking to demote the markedness constraint below the faithfulness constraint.

Our characterization also reveals some important differences between these error patterns, which allow us to predict when overgeneralization is and is not expected to occur. On the one hand, overgeneralization errors are seen to be an

---

account of the acquisition facts. Our schema is, however, capable of accounting for allophonic phenomena in developing and fully developed languages. It remains to be determined whether all allophonic phenomena are more properly characterized in terms of complementary constraints as in our schema.

expected, possibly unavoidable, consequence of interacting markedness constraints. On the other hand, complementary errors are seen to be resistant to overgeneralization because their associated markedness constraints cannot interact.

As a final comment, I would like to suggest that findings of this sort offer promise that optimality theoretic accounts of other phenomena will yield additional new insights about acquisition. They also serve to underscore the role that acquisition can play in the evaluation of contemporary phonological theories.

References

Barlow, Jessica A. and Gierut, Judith A. 1999. Optimality theory in phonological acquisition. Journal of Speech, Language, and Hearing Research, 42.1482-98.

Bernhardt, Barbara H. and Stemberger, Joseph P. 1998. Handbook of phonological development from the perspective of constraint-based non-linear phonology. San Diego, CA: Academic Press.

Camarata, Stephen M. and Gandour, Jack. 1984. On describing idiosyncratic phonologic systems. Journal of Speech and Hearing Disorders, 49.262-66.

Chomsky, Noam and Halle, Morris. 1968. The sound pattern of English. New York: Harper & Row.

Dinnsen, Daniel A. 1996. Context-sensitive underspecification and the acquisition of phonemic contrasts. Journal of Child Language, 23.57-79.

Dinnsen, Daniel A. 1999. Some empirical and theoretical issues in disordered child phonology. Handbook of child language acquisition, ed. by W. C. Ritchie and T. K. Bhatia, 647-74. New York: Academic Press.

Dinnsen, Daniel A. and Chin, Steven B. 1995. On the natural domain of phonological disorders. Phonological acquisition and phonological theory, ed. by J. Archibald, 135-50. Hillsdale, NJ: Erlbaum.

Dinnsen, Daniel A., O'Connor, Kathleen M. and Gierut, Judith A. 2001. The puzzle-puddle-pickle problem and the Duke-of-York gambit in acquisition. Journal of Linguistics, 37. 507-25.

Donegan, Patricia J. and Stampe, David. 1979. The study of natural phonology. Current approaches to phonological theory, ed. by D. A. Dinnsen, 126-73. Bloomington, IN: Indiana University Press.

Eckman, Fred R. and Iverson, Gregory K. 1997. Structure preservation in interlanguage phonology. Focus on phonological acquisition, ed. by S. J. Hannahs and M. Young-Scholten, 183-208. Amsterdam: John Benjamins.

Fey, Mark E. 1989. Describing developing phonological systems: A response to Gierut. Applied Psycholinguistics, 10.455-67.

Gierut, Judith A. 1986. Sound change: A phonemic split in a misarticulating child. Applied Psycholinguistics, 7.57-68.

Gierut, Judith. A. 1989. Developing descriptions of phonological systems: A surrebuttal. Applied Psycholinguistics, 10.469-73.

Gierut, Judith A. and Champion, Annette H. 2000. Ingressive substitutions: Typical or atypical phonological pattern? Clinical Linguistics & Phonetics, 14.603-17.

Goldman, Ronald and Fristoe, Macalyne. 1986. Goldman-Fristoe Test of Articulation. Circles Pines, MN: American Guidance Service.

Hardy, Jane E. 1993. Phonological learning and retention in second language acquisition. Confluence: Linguistics, L2 acquisition and speech pathology, ed. by F.R. Eckman, 235-48. Amsterdam: John Benjamins.
Ingram, David. 1989. Phonological disability in children. London: Cole and Whurr.
Kager, Rene. 1999. Optimality theory. Cambridge: Cambridge University Press.
Lado, Robert. 1957. Linguistics across cultures. Ann Arbor: University of Michigan Press.
Leonard, Lawrence B. and Brown, Barbara L. 1984. Nature and boundaries of phonologic categories: A case study of an unusual phonologic pattern in a language-impaired child. Journal of Speech and Hearing Disorders, 49.419-28.
Macken, Marlys A. 1980. The child's lexical representation: The 'puzzle-puddle-pickle' evidence. Journal of Linguistics, 16.1-17.
Macken, Marlys. A. and Ferguson, Charles A. 1983. Cognitive aspects of phonological development: Model, evidence and issues. Children's language, ed. by K. E. Nelson, 255-82. Hillsdale, NJ: Erlbaum.
McCarthy, John J. and Prince, Alan S. 1995. Faithfulness and reduplicative identity. University of Massachusetts Occasional Papers 18: Papers in optimality theory, ed. by J. N. Beckman, L. W. Dickey and S. Urbanczyk, 249-384. Amherst, MA: GLSA.
Menn, Lise. 1983. Development of articulatory, phonetic, and phonological capabilities. Language production: development, writing and other language processes, ed. by B. Butterworth, 3-50. New York: Academic Press.
Prince, Alan and Smolensky, Paul. 1993. Optimality theory: Constraint interaction in generative grammar. New Brunswick, NJ: Rutgers Center for Cognitive Science, Rutgers University.
Sheldon, Amy and Strange, Winifred. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. Applied Psycholinguistics, 3.243-61.
Smith, Neil V. 1973. The acquisition of phonology: A case study. Cambridge, UK: Cambridge University Press.
Smolensky, Paul. 1996a. The initial state and 'richness of the base' in optimality theory (Technical Report No. JHU-CogSci-96-4). Baltimore, MD: Department of Cognitive Science, Johns Hopkins University.
Smolensky, Paul. 1996b. On the comprehension/production dilemma in child language. Linguistic Inquiry, 27.720-31.
Tesar, Bruce and Smolensky, Paul. 1998. Learnability in optimality theory. Linguistic Inquiry, 29.229-68.
Williams, A. Lynn and Dinnsen, Daniel A. 1987. A problem of allophonic variation in a speech disordered child. Innovations in Linguistic Education, 5.85-90.