LANGUAGE ACQUISITION, *8*(4), 321–361

On the Role of Sympathy in Acquisition

Daniel A. Dinnsen, Laura W. McGarrity,
Kathleen M. O'Connor, and Kimberly A. B. Swanson
Indiana University

Evidence from young children's early phonological development is brought to bear on the evaluation of a newly proposed type of correspondence relation within optimality theory (McCarthy and Prince (1995), Prince and Smolensky (1993)) namely *sympathy*. Sympathy has been advanced to account for certain opacity effects in fully developed languages. Given the claims of the theory, comparable opacity effects are expected to occur in the course of acquisition. Toward this end, different interactions of two common phenomena—that is, final consonant omission and vowel lengthening before voiced consonants—are examined with a focus on a case study of 2 young children with phonological delays in their acquisition of English. We argue that at least some developmental opacity effects support sympathy and that such effects naturally emerge in the course of development from the harmonic ranking of sympathy over input–output faithfulness and the incremental demotion of markedness constraints.

1. INTRODUCTION

It is generally acknowledged that independent phonological phenomena can interact in ways that may obscure generalizations (or render them opaque). In rule-based derivational theories of phonology (e.g., Chomsky and Halle (1968)), these opacity effects obtain largely from intermediate levels of representation and characteristic rule ordering relations among two or more rules. For example, a counterbleeding order between two rules results in phonetic forms in which one of the rules has applied but it appears that it should not have (given that the conditioning environment is no longer evident). A counterfeeding order between two

Requests for reprints should be sent to Daniel A. Dinnsen, Department of Linguistics, Indiana University, Memorial Hall East 332, 1021 East Third Street, Bloomington, IN 47405–7005. E-mail: dinnsen@indiana.edu

rules results in forms in which it appears that one of the rules should have applied but did not. These same opacity effects pose a different set of challenges for more recent constraint-based theories such as optimality theory (OT), in which it is hypothesized that there are no intermediate levels of representation, no rules, and thus, no rule ordering (e.g., McCarthy and Prince (1995), Prince and Smolensky (1993)). The theory in its most basic form would predict that phonetic forms be transparent (or fully consistent with surface true generalizations). The theory does, however, acknowledge the existence of opacity effects and has attempted to account for many of them, including, for example, those evident in reduplication (McCarthy and Prince (1995)), truncation (Benua (1995)), chain shifts (Kirchner (1996)), and nonderived environment blocking (Łubowicz (1998)). In an effort to deal with certain other opacity effects in fully developed languages, McCarthy (1999b; 1999c) offered a further modification to the theory by introducing a novel type of correspondence relation—namely, *sympathy.* Sympathy differs from other correspondence relations in important respects and has many implications for theory and acquisition, as we see next.

First, by way of background, in OT a ranked set of universal constraints evaluates the set of all possible output candidates for any given input representation and selects one output as optimal by virtue of its better satisfying the constraint hierarchy. The constraints have been understood to fall into at least two general types: faithfulness and markedness (the latter also referred to as structural or well-formedness constraints). The faithfulness constraints demand an identical match or correspondence between two strings. In some instances, the two strings refer to an input representation and an output representation and are related by an input–output (IO) faithfulness constraint. If an IO faithfulness constraint is not dominated by a competing markedness constraint, underlying contrasts are preserved in the output. The analog in derivational terms would be the absence of a phonological rule. A correspondence relation can also hold between two strings in which both are potential output candidates (output–output or OO faithfulness). In such cases, there is a correspondence between, for example, the base form of a word and a morphologically related form of the word (truncation or reduplication). The markedness constraints tend to be antagonistic to faithfulness, favoring outputs that are structurally unmarked. When markedness constraints are highly ranked, they achieve many of the same effects associated with phonological rules in derivational theories. The undominated ranking of either type of constraint results in the optimal output being relatively transparent.

Sympathy extends the theory by allowing an opaque output to be selected as optimal on the basis of its correspondence with a “failed” output candidate. Briefly, the relevant failed candidate is considered the object of sympathy and is termed the *flower candidate.* The candidate that serves as the flower candidate is determined by some other constraint. McCarthy (1999b; 1999c) argued that the flower candidate is identified from a set of candidates that complies with and is defined by a low-ranked faithfulness constraint (i.e., the selector).[1] By virtue of the selector constraint being low ranked (i.e., dominated by a markedness constraint), many candidates that are consistent with the selector are failed candidates. The candidate of that set that is the most harmonic relative to the other constraints serves as the flower candidate. The sympathy constraint preserves some property of the flower candidate in another output candidate. If the flower candidate and the input representation happened to be different, and if the sympathy constraint outranked the conventional IO faithfulness constraints, the optimal output candidate would be opaque, more closely resembling the flower candidate than the input representation.

Sympathy as a theoretical construct has received further support from its successful accounts of various phenomena in fully developed languages (e.g., Davis (1997), de Lacy (1998), Itô and Mester (1997), Koontz-Garboden (2000), Lee (1999), McGarrity (1999), Walker (1999)).[2] Nevertheless, many questions are raised by sympathy, and a broader range of test implications must be sought. Some of these as yet unanswered questions are, Do comparable opacity effects occur in developing systems, and how do they emerge in a child's grammar? What role can sympathy play in accounting for those opacity effects? Are sympathy constraints universal, and do they participate in any universal or initial state (default) ranking relations with other constraints? Can different constraints serve as selectors, identifying multiple competing flower candidates to yield a virtual bouquet of flower candidates? If so, how can the one empirically correct flower candidate be selected? This latter question might be dubbed the *bouquet problem.*

Answers to such questions and a further test of sympathy might reasonably reside in acquisition phenomena from the developing systems of young children. One reason for this is that early stages of development with their many production errors (relative to the target system) have generally been characterized by a number of markedness constraints outranking faithfulness constraints (e.g., Gnanadesikan (1996), Smolensky (1996b); cf. Dinnsen and Barlow (1998a), Hale and Reiss (1998)). This means that there should be many low-ranked faithfulness constraints that could potentially serve as selectors, identifying a flower candidate. Properties of that flower candidate could in turn be preserved by some sympathy constraint. In short, we should expect opacity effects in developing systems, and they should be amenable to sympathy.

The purpose of this article is to assess the role of sympathy in acquisition. We hope to show that at least some of the opacity effects that do occur in developing systems are quite tractable with sympathy and, more important, that an appeal to

---

[1]For arguments that the selector can also be a markedness constraint, see Itô and Mester (1997), de Lacy (1998), and Walker (1999).

[2]For an alternative approach to some of these opacity effects within OT, see Sprouse (1998), Goldrick and Smolensky (1999), and Itô and Mester (1999).

sympathy offers some new insights for both acquisition and phonological theory. For acquisition in particular, sympathy is shown to contribute to a characterization of individual differences and to a plausible account of the emergence of opacity in the transition from early stages of development to an end state. The contribution to phonological theory is the added support that accrues to sympathy from the venue of acquisition and the insight that it affords into the ranking relation between sympathy and other types of constraints. Finally, a preliminary and partial solution is offered for resolving the bouquet problem. The focus is on the interaction of two common phenomena—vowel lengthening before voiced consonants and the error pattern of final consonant omission—in the speech of two children with phonological delays in their acquisition of English.

2. VOWEL LENGTHENING AND FINAL CONSONANT OMISSION

In this section, we first consider the interaction of the two phenomena—vowel lengthening and final consonant omission—for a child who evidenced an opaque interaction. A sketch of a derivational account is formulated and compared to some likely optimality accounts that exclude sympathy but that are shown to fail in various respects. An optimality account employing sympathy is then formulated and evaluated against the facts at hand. The evaluation is extended to the characterization of individual differences and development through consideration of a second child who exhibited a transparent interaction. Sympathy, we argue, provides a viable optimality theoretic account of the facts for both children while also preserving a high degree of continuity across stages of development. As a consequence of our developmental considerations, we advance a harmonic (universal) ranking relation between sympathy and IO faithfulness.

2.1. Background and Facts

The phenomenon of vowel lengthening before voiced consonants is a near universal, being widely attested in the languages of the world (e.g., Chen (1970)) as well as in children's early speech (e.g., Raphael, Dorman, and Geffner (1980), Weismer (1984)). Despite the ubiquity of this phenomenon, there is surprisingly little agreement on its characterization. First, although it is often assumed that the lengthening effect is the result of some automatic phonetic process, this must be reconciled against the fact that some languages, such as Arabic (Port, Al-Ani, and Maeda (1981)) and Polish and Czech (Keating (1979)), do not evidence the effect. As is seen next, this problem is further complicated by those cases in which lengthening occurs even when the presumed phonetic conditioning is absent. Even if lengthening were understood to be language specific, the controversy would likely remain over its phonetic–phonological character (e.g., Keating

(1985)). There are also questions about how vowel length is to be represented in terms of a categorical feature [long] (e.g., Pyle (1971)), a gradient feature (e.g., Dinnsen and Charles-Luce (1984)), or structurally in terms of vowels being associated with a moraic or skeletal tier (e.g., Clements (1986), Hubbard (1995)). Although these issues are beyond the scope of this article, the relevant generalization remains that vowels lengthen before voiced consonants, at least in English, for both adults and many children. We assume for expository purposes that a feature [long] is available for the expression of generalizations about vowel length, but this is not intended as a formal proposal. Our conclusions should apply equally to alternative characterizations of length.

The phenomenon of vowel lengthening can and does interact with a cross-linguistically common error pattern in phonological development (normal or disordered)—namely, the omission of final consonants (Locke (1983)). For children with normal development acquiring English, Smit (1993) found that all age groups between 2;0 and 8;0 exhibited final consonant omissions with varying degrees of frequency. For children with a phonological delay or disorder, the error pattern may persist somewhat longer or occur with greater regularity, or both (Stoel-Gammon and Dunn (1985)). The potential for an interaction between final consonant omission and vowel lengthening arises because final consonants can constitute part of the conditioning environment for lengthening. In the course of acquisition, these phenomena can co-occur and reveal a crucial interaction. For example, in an instrumental study of the speech of three children with phonological delays who omitted final obstruents, Weismer, Dinnsen, and Elbert (1981) provided measures of vowel length durations before omitted word-final voiced and voiceless obstruents. It was found that two of the three children maintained a statistically significant vowel length distinction such that vowels were long before (omitted) voiced obstruents and short elsewhere. The tokens in (1) illustrate this result for one of the children of that study, Child A (age 7;2).[3]

(1) Child A (age 7;2)

a.	kæː	‘cab’	ka	‘cop’
	kɪː	‘kid’	pæ	‘pat’
	dɔː	‘dog’	dʌ	‘duck’
b.	kæbi	‘cabby’	kapoʊ	‘copper’
	kidoʊ	‘kidder’	pæti	‘patty’
	dɔgi	‘doggie’	dʌki	‘ducky’

Word-final obstruents (but not nasals) were omitted with a high degree of regularity in this child’s speech, especially in conversational samples. In the more struc-

[3]Similar results obtained for Child B (age 7;6), although with a lower percentage of omission errors. However, if not omitted, word-final obstruents were replaced by a glottal stop. The target voice contrast was nonetheless preserved in the preceding vowel length, as was the case for Child A. The lengthening effect was rendered opaque as a result of the merger of obstruents to a glottal stop.

tured elicitation task, word-final obstruents were omitted in 80% of the tokens that were produced in isolation or at the end of an utterance. The remainder of the forms in that same context either were produced with a final glottal stop or did not receive unanimous agreement from the three judges that the obstruent was omitted.[4] Although this child fell within normal age limits for the occurrence of this error pattern (Smit (1993)), the number of other errors and the absence of any identifiable organic problems suggested that the child be classified as phonologically delayed with a functional speech disorder.

In addition to the omission of final obstruents, the transcriptions in (1a) show corresponding vowel length differences in accord with the voicing of the omitted obstruent. The forms in (1b) establish that there was a consonantal alternation and show that the stem-final obstruents did occur when not in word-final position.[5] The transcriptions of these latter forms do not indicate vowel length because no measurements were made for these words. However, because vowel length is one of the primary perceptual cues for postvocalic voicing differences in obstruents, it is assumed here, consistent with the other cited instrumental studies, that the transcription of voicing reflects the fact that vowels were transparently (allophonically) long before the occurring voiced obstruents and short before voiceless.

2.2. A Sketch of a Derivational Account

In derivational terms, these facts could readily be accounted for by postulating underlying representations that are essentially adult-like—that is, with short vowels and stem-final voiced and voiceless obstruents.[6] A rule of vowel lengthening would be necessary to convert those short vowels to their predictably long counterparts before voiced consonants. An additional rule of final consonant deletion would account for the omission of final obstruents and the alternation between

[4]In sentence-medial contexts, word-final obstruents were judged to be omitted less frequently (in approximately 40% of the tokens). A possible explanation for the lower rate of omissions in this context is that the obstruents were not in a true word-final context. That is, they were produced in the middle of what might be considered a compound word. It is also interesting that vowel length was not differentiated in this context, probably due to the natural compression that occurs in longer utterances. The failure of that context to sustain vowel length differences may have prevented obstruent omissions on recoverability grounds. If obstruents had been omitted without preserving a vowel length distinction, the voice contrast in obstruents would not have been recoverable.

[5]Independent of the morphological relatedness of the word pairs in this data set, they do at least serve to illustrate distributional properties of vowel length and postvocalic obstruents.

[6]We take these assumptions to be relatively uncontroversial in derivational frameworks in which simplicity considerations apply. Thus, any alternative account that might posit underlying long and short vowels for different words is judged less desirable because it misses the generalization about vowel length before the occurring obstruents and results in an increase in the number of vowel phonemes. It would also be forced to reject the widely held assumption that children's underlying representations are adult-like.

word-final null and the occurring word-medial obstruents. As illustrated in (2), these rules would interact such that the lengthening rule must apply before deletion in a counterbleeding relation. If deletion had applied first, lengthening would have been bled, resulting in a transparent but unattested output. For words ending with an underlyingly voiced obstruent, the actual output is opaque because a long vowel appears in a context other than before a voiced consonant.

(2) A derivational account (counterbleeding order)

Vowel lengthening (VL): V → [long]/__[consonantal, voice]

Final consonant deletion (FCD): [-sonorant] → Ø/__#

	/kæb/	'cab'	/pæt/	'pat'
VL	kæːb		-----	
FCD	kæː		pæ	
	[kæː]		[pæ]	

We return to a consideration of derivational accounts of these and related facts after formulating a viable optimality account of the same facts.

2.3. A Conventional Optimality Account

An optimality theoretic account of these facts implicates at least the constraints in (3).

(3) Some relevant constraints and a preliminary ranking

a. Markedness constraints

LENGTHEN: Avoid short vowels before voiced consonants; avoid long vowels elsewhere.

NOCODA: Avoid obstruents in codas.

b. Faithfulness constraints

ID[weight]: The length (or weight) of corresponding vowels in the input and output should be identical.

MAX: Every input segment has a corresponding output segment (no deletion).

Ranking: NOCODA, LENGTHEN >> ID[weight], MAX

In the ranking statement in (3), the notation >> is to be interpreted as "outranks." A comma between two constraints indicates that they are unranked relative to one another or that the ranking is indeterminate.

LENGTHEN is the markedness constraint that would account for the general lengthening effect by favoring long vowels before voiced consonants. In all other contexts, long vowels would be disfavored. LENGTHEN is an abbreviation for several constraints affecting vowel length and is not intended to be exhaustive. That is, there may also be other contexts in which long vowels are favored—for example, in phrase-final position, or in stressed (vs. unstressed) syllables. The point is, however, that vowel length is not contrastive in any of these contexts. One example of a violation for LENGTHEN would be any candidate with a short vowel before a voiced consonant. As formulated, the constraint also acknowledges the relatively marked character of long vowels in certain other contexts. For example, a violation would be incurred by any candidate with a long vowel before a voiceless consonant. The undominated ranking of this constraint would result in vowel length being noncontrastive in any context. The other markedness constraint, NOCODA, accounts for the omission of final obstruents by favoring output candidates without a coda consonant. It is this same highly ranked constraint that accounts for the cross-linguistically common prohibition against coda consonants.

The relevant antagonistic faithfulness constraints assess violations for a candidate's failure to preserve properties of the input representation. MAX is the standard faithfulness constraint militating against deletion. This constraint would be violated by any candidate that failed to preserve an input segment. The faithfulness constraint that is responsible for preserving a vowel's input length or weight is ID[weight].[7] This constraint would be violated by any candidate that differed from the input representation in terms of vowel length. If a short vowel were posited underlyingly (as in our derivational account earlier), the occurrence of a corresponding long vowel in that word would violate this constraint. Similarly, if a long vowel were posited in the input representation of a word, any corresponding output candidate with a short vowel would violate the constraint. Given that some of the occurring forms in this child's speech tolerate violations of MAX and that they largely comply with LENGTHEN (no matter what might be assumed about input vowel length), we begin with the assumption that the two faithfulness constraints are low ranked.

The tableau in (4) considers four likely output candidates for the word *cab*. Consistent with conventions in OT, competing output candidates are listed down the left side of the tableau, and constraints are arranged across the top, reflecting their relative ranking. Higher ranked constraints are to the left, and lower ranked constraints to the right. A solid vertical line separates crucially ranked constraints, whereas a dotted vertical line indicates that the two constraints are unranked relative to one another. The input representation is given in the upper left-hand corner. Violations of constraints are noted by an asterisk (*), with fatal violations being marked by an exclamation point (!). Fatal violations eliminate a candidate from consideration. The optimal or winning candidate as predicted by the constraint ranking is noted by the manual indicator (☞).

[7]We set aside here the issue of whether faithfulness is a bidirectional or unidirectional correspondence relation (cf. MAX-μ and/or DEP-μ). If MAX-μ and DEP-μ were equally ranked, they would have the same effect as ID[weight].

(4) Incorrect prediction of transparent output

'cab' /kæb/	NoCoda	Lengthen	Max	ID[weight]
a. kæb	*!	*		
b. kæ:b	*!			*
c. ☞ kæ			*	
d. → kæ:		*!	*	*

Assuming the ranking in (3), the first two candidates with a final consonant can be eliminated due to their violation of undominated NoCoda. The faithful candidate (a) also incurs a violation of the markedness constraint Lengthen, which demands that vowels be long before voiced consonants, whereas candidate (b) incurs a violation of ID[weight]. Of the two remaining candidates without the final consonant, candidate (d) with the long vowel should win (as indicated by the arrow), because that is Child A's actual output, but is eliminated because of its fatal violation of Lengthen. The transparent output candidate (c) with a short vowel and no final consonant is incorrectly predicted to be the optimal candidate.

Although this ranking fails to predict the actually occurring form, there is in fact no available ranking of these constraints that is capable of making the empirically correct prediction. One reason for this is that candidate (c) has a subset of the violations of candidate (d). More specifically and as shown in (5), given the assumption of short vowels in input representations and the alternate ranking Max, ID[weight] >> NoCoda, Lengthen, the faithful candidate (a) would incorrectly be selected as optimal.

(5) Faithfulness over markedness

'cab' /kæb/	Max	ID[weight]	NoCoda	Lengthen
a. ☞ kæb			*	*
b. kæ:b		*!	*	
c. kæ	*!			
d. → kæ:	*!	*		*

Similarly, as shown in (6), the alternate ranking Max, Lengthen >> NoCoda, ID[weight] would incorrectly select candidate (b). This is in fact the ranking that would ultimately be required for adult English.

(6) Ranking for Adult English

'cab' /kæb/	MAX	LENGTHEN	NOCODA	ID[weight]
a. kæb		*!	*	
b. ☞ kæ:b			*	*
c. kæ	*!			
d. → kæ:	*!	*		*

As shown earlier, the alternate ranking in tableau (4) incorrectly selects candidate (c) with a short vowel and no final obstruent. This inability to predict opacity in those cases in which it is needed is precisely the problem confronting conventional OT.

2.4. A Brute Force Account

The difficulties with the preceding optimality account force us to consider an alternative that adopts a different set of assumptions about the underlying input representations. That is, rather than assuming that the affected vowels are underlyingly short, it might be assumed that there is an underlying vowel length distinction with some vowels marked as long in input representations and others marked as short. This alternative might seem plausible given that both long and short vowels did occur in the child's speech and in the primary linguistic data to which the child would have been exposed. However, vowel length is generally assumed to be noncontrastive in the target system, making it unclear what fact would motivate the child to rank ID[weight] over LENGTHEN to ensure a length distinction. Nevertheless, vowel length is superficially contrastive in the child's speech, at least in a limited set of contexts. To ensure that the putative underlying length distinction is realized in the child's output where posited, a different ranking of the constraints would be necessary with ID[weight] ranked above LENGTHEN. Thus, if long vowels were posited before all voiced consonants and short vowels before all voiceless, then the empirically correct vowel length can be ensured, even when the final obstruent is omitted in the output. The tableaux in (7) and (8) illustrate this point for two representative words differing in the voicing of the final obstruent.

(7) Input long vowels before voiced consonants

'cab' /kæ:b/	NOCODA	ID[weight]	MAX	LENGTHEN
a. kæb	*!	*		*
b. kæ:b	*!			
c. kæ		*!	*	
d. ☞ kæ:			*	*

(8) Input short vowels before voiceless consonants

'pat' /pæt/	NoCoda	ID[weight]	Max	Lengthen
a. pæt	*!			
b. pæːt	*!	*		*
c. ☞ pæ			*	
d. pæː		*!	*	*

Candidates (a) and (b) in tableaux (7) and (8) are eliminated as a result of their violations of undominated NoCoda. The choice between candidates (c) and (d) is made by high-ranked ID[weight]. In tableau (7), candidate (c) with a short vowel is eliminated for its failure to preserve input vowel length. Thus, candidate (d) with a long vowel wins even though it violates lower ranked Lengthen. In tableau (8), candidate (d) with a long vowel is eliminated for its failure to preserve the input short vowel length. Thus, candidate (c) with a short vowel wins.

Although this alternative with its associated constraint ranking and assumptions of distinctive vowel length would seem to account for the facts at hand, it is limited in several respects. First, if vowel length were truly contrastive, the long and short vowel phonemes would be widely distributed. Instead, these vowel phonemes are required to have a defective distribution that is largely complementary. The fact is that the superficial length contrast occurs only in the context before omitted obstruents. Otherwise, the distribution of long and short vowels is entirely predictable. The account must treat as accidental the complementary distribution of vowel length in the transparent cases—that is, when the conditioning consonant does occur phonetically (1b). The only reason by this account that vowels are long before voiced consonants and short before voiceless is because high-ranked faithfulness demands that they be realized as they are underlyingly represented. Similarly, the lower ranking of Lengthen predicts that violations of that markedness constraint should be tolerated, yet this account fails to explain the nonoccurrence of short vowels before voiced consonants and long vowels before voiceless consonants in phonetic outputs. It must also treat as accidental the absence of a vowel length contrast in words that do not exhibit a consonantal alternation. For example, a word such as *do* is realized without a postvocalic consonant even in its derived form (*doing*). If the occurrence of an underlying postvocalic consonant were irrelevant to the conditioning of preceding vowel length, we might expect to see a vowel length contrast in the base form of such words. The singular value of this account centers on its ability to yield a long vowel before an omitted voiced obstruent. This is achieved essentially by brute force—namely, by relying on the substance of underlying representations with a defective distribution.[8]

[8]An anonymous reviewer suggested that the brute force account would predict that the child would

OT has employed two principles relating to the substance of input representations, both of which are relevant to the issue here. One principle, richness of the base (Prince and Smolensky (1993), Smolensky (1996a)), maintains that the set of input representations is universal. This means that whatever is a possible underlying representation in one language should be a possible underlying representation in all other languages. Languages are thus the same in this regard. Any account of a particular language must allow for the full set of possible underlying representations. There can be no limitations or restrictions on the underlying representations within a particular language. The fact that not all underlying contrasts are realized in a language must be handled by the constraint hierarchy. Specifically, it is left to certain high-ranked markedness constraints to eliminate candidates with those features that do not occur within that particular language.

This has immediate consequences for any optimality theoretic account of the facts being considered here. That is, although the brute force account might seem at first to be consistent with richness of the base in that it allowed both long and short vowels in input representations, it actually violates the principle by restricting the distribution of these vowels in input representations. Recall that long vowels had to be restricted from occurring before voiceless consonants, and short vowels had to be restricted from occurring before voiced consonants at the underlying level (and at the phonetic level). The nonoccurrence of a vowel length contrast in nonalternating words (e.g., *do, doing*) would require a similar restriction on input representations. If the distribution of these vowels had not been so restricted in the input representations, the high-ranked faithfulness constraint ID[weight] would have generated precisely what did not occur and was presumably impermissible—namely, outputs with long vowels before voiceless consonants, short vowels before voiced consonants, and a vowel length contrast in those words with no consonantal alternation. For any account to be fully consistent with richness of the base, it must allow for the possibility that a long or short vowel could underlie any vowel in any context to yield a permissible output. The brute force account would thus be ruled out by richness of the base.

The other principle relating to the substance of underlying representations is lexicon optimization (e.g., Itô, Mester, and Padgett (1995), Prince and Smolensky (1993)). This principle provides a means for selecting the one presumably correct underlying representation for a word from among possible alternative underlying representations of that same word. This principle has been invoked to address abstractness and learnability concerns. The issue is especially relevant for nonalternating forms governed by high-ranked markedness constraints. Consider, for

---

produce a long vowel before a voiceless consonant if exposed to such a sequence in his primary linguistic data. It is, however, difficult to evaluate this prediction given the absence of such sequences in English. In any event, even if such sequences were to occur and the child were to recognize them as such, this should result in the demotion of LENGTHEN below ID[weight], wiping out any presumed differences in predictions.

example, the case in which a child might systematically and exclusively produce long vowels before voiced consonants (e.g., [dɔːgi] 'doggie'). In such a case, the markedness constraint LENGTHEN would be claimed to outrank the faithfulness constraint ID[weight]. As a result of this ranking and consistent with richness of the base, the correct output with a long vowel would obtain no matter what is assumed about input vowel length. Nevertheless, there would be two possible input representations for this one winning output candidate. In the case in which the input includes a long vowel, the winning output candidate would incur no violations of either constraint. On the other hand, and in the instance in which the other possible input representation includes a short vowel, the winning output candidate with a long vowel would incur a violation of the lower ranked faithfulness constraint ID[weight]. Although the violation of this lower ranked faithfulness constraint would not affect the output, it is presumed to bear on the selection (learning) of an optimal input. Lexicon optimization would select the input that is the more harmonic of the two—that is, the one that incurs the fewest violations. In this case, then, for the word [dɔːgi], the optimal input would be claimed to include a long vowel.

Although it might seem that lexicon optimization and the brute force account concur on the appropriateness of postulating long and short vowels in at least certain cases for Child A, any such conclusion would at best be misleading or irrelevant. Recall that the brute force account had to rank ID[weight] above LENGTHEN. As a result, any winning output candidate would have one and only one possible input representation (without ever having to appeal to lexicon optimization). That is, every occurring long vowel would have corresponded exclusively to an input long vowel, and every occurring short vowel would have corresponded exclusively to an input short vowel. Any issue of selecting from among competing input representations for a winning output candidate would have been precluded by high-ranked ID[weight] and the defective distribution of long and short vowel phonemes. Because the brute force account must violate richness of the base in its account of the nonoccurrence of long vowels before voiceless consonants and of short vowels before voiced consonants, the issue of lexicon optimization is again precluded. The brute force account effectively prevents lexicon optimization from contributing to the substance of these underlying representations.

The conclusion that emerges from a consideration of these two principles is that the brute force account accrues no support for any of its claims from either richness of the base or lexicon optimization. In fact, the brute force account is incompatible with richness of the base. Whether these empirical and theoretical limitations outweigh any of the presumed benefits can only be determined by a consideration of some other viable account of the same facts. We see in what follows that sympathy offers an account of these and other facts in a way that complies with richness of the base and lexicon optimization and better conforms to other widely held assumptions about development.

2.5. The Sympathy Account

We return to our original optimality account and reconsider the troublesome fact about the occurrence of long vowels before omitted voiced obstruents. The tableau in (4) is most relevant and is carried over with certain amendments in (10). It can be observed that the empirically correct opaque output candidate (d) most closely resembles candidate (b) in terms of vowel length. This is suggestive of a sympathy constraint, ❀SYM in (9), which attempts to preserve a correspondence in vowel length between certain output candidates.

(9) Sympathy constraint
❀SYM: The length (or weight) of corresponding vowels in the flower candidate and an output candidate should be identical.

(10) High-ranked sympathy correctly predicts an opaque output[9]:

'cab' /kæb/	NOCODA	❀SYM	LENGTHEN	MAX	ID[weight]
a. kæb	*!	*	*		
b. ❀ kæ:b	*!				*
c. kæ		*!		*	
d. ☞ kæ:			*	*	*

Although candidate (b) is a failed candidate (due to its violation of NOCODA), it must nonetheless serve as the object of sympathy in this case and is thus designated as the flower candidate (indicated by ❀). According to McCarthy (1999b; 1999c), the general principles for identifying a flower candidate call for a low-ranked faithfulness constraint to serve as the selector, in this case MAX.[10] As such, MAX defines a set of candidates that complies with the constraint. That set includes all candidates with a final consonant—namely, candidates (a) and (b). These candidates are acknowledged to be failed candidates due to their violation of high-ranking NOCODA. Of those failed candidates, candidate (b) is the more harmonic with respect to the rest of the constraint hierarchy, only violating low-ranked ID[weight] if input short vowels were assumed. Candidate (a) violates the high-ranked markedness constraint LENGTHEN and thus could not be the object of sympathy conferred by the selector MAX. Sympathy constraints are assumed to be invisible in the selection of the flower candidate. Thus, although candidate (a) would violate the sympathy constraint, the violation does not count for the purposes

[9]Although the sympathy constraint must outrank LENGTHEN, the ranking relation between NOCODA and LENGTHEN is indeterminate.

[10]There are a number of other low-ranked faithfulness constraints that might also serve as selectors, in particular ID[weight], which would allow for multiple flower candidates. Aside from ID[weight], none of the other low-ranked faithfulness constraints should affect the outcome. The apparent problem of selecting the one effective flower candidate is dealt with in section 3.1.

of selecting the flower candidate. By ranking ❀SYM above the IO faithfulness constraints, it is claimed that it is more important to preserve the vowel length of the flower candidate than it is to preserve the vowel length of the input representation (if underlying short vowels were assumed). Similarly, by ranking ❀SYM above LENGTHEN, it is claimed that it is more important to be faithful to the vowel length of the flower candidate than it is to comply with LENGTHEN, even if a long vowel results. This is precisely what is called for in this case. The tableau in (10) can thus account for the realization of *cab* by employing high-ranked sympathy.

Many of the points from our original optimality account have been retained here. The first two candidates with a final consonant are eliminated by the high-ranked markedness constraint NOCODA. The choice is then between the two candidates without the final consonant. Although candidate (c) with the short vowel is faithful to input vowel length under one assumption and complies with LENGTHEN, it incurs a violation of ❀SYM for its failure to preserve the vowel length of the flower candidate. Candidate (d) with the long vowel violates LENGTHEN and low-ranked ID[weight] but complies with ❀SYM. If ❀SYM and LENGTHEN were equally ranked, a tie would result, passing the choice to the lower ranked faithfulness constraints that would incorrectly select candidate (c). Therefore, ❀SYM must outrank LENGTHEN to eliminate candidate (c). The opaque candidate is the most harmonic candidate given this ranking and is thus selected as optimal.

Assuming the same constraints and constraint ranking, the empirically correct transparent output is predicted for other target words that end with a voiceless consonant such as *pat,* as illustrated in the tableau in (11). First, the flower candidate identified by low-ranked MAX is the faithful candidate (a) because that candidate complies with LENGTHEN and ID[weight], whereas candidate (b), which complies with MAX, does not. As for the ultimate evaluation of candidates, the two candidates (a) and (b) with a final consonant are eliminated by high-ranked NOCODA. The competition is thus between the two candidates without a final consonant. Candidate (d) with a long vowel incurs a fatal violation of ❀SYM because it fails to preserve the length of the flower candidate (which happens to be the same as the input vowel length). That candidate also violates LENGTHEN due to the occurrence of a long vowel in a context other than before a voiced consonant. Candidate (c) with a short vowel complies with all constraints, except low-ranked MAX, and thus wins out.

(11) Sympathy and correct transparent output

'pat' /pæt/	NOCODA	❀SYM	LENGTHEN	MAX	ID[weight]
a. ❀ pæt	*!				
b. pæːt	*!	*	*		*
c. ☞ pæ				*	
d. pæː		*!	*	*	*

The sympathy solution presented here accounts for the critical fact about the occurrence of long vowels before omitted voiced obstruents, also avoiding the shortcomings of the brute force account considered earlier. One of those shortcomings was its need to assume an underlying vowel length distinction in input representations. The sympathy account yields the correct results no matter what is underlyingly assumed about vowel length. Our discussion has been formulated assuming underlying short vowels, but the same results obtain if long vowels or a vowel length distinction were also assumed in inputs. Thus, even if a long vowel were posited in the input for *cab* as in (12) (cf. the tableau in (10) with an alternative input short vowel), the same flower candidate would be chosen and the empirically correct output with a long vowel would result due to the ranking of ❀SYM over LENGTHEN.

(12) Input long vowels before voiced consonants

'cab' /kæ:b/	NOCODA	❀SYM	LENGTHEN	MAX	ID[weight]
a. kæb	*!	*	*		*
b. ❀ kæ:b	*!				
c. kæ		*!		*	*
d. ☞ kæ:			*	*	

In addition, by ranking the markedness constraint LENGTHEN above the faithfulness constraint ID[weight], any assumptions about input vowel length will be sacrificed in favor of markedness. Thus, even if an input long vowel were posited before a voiceless obstruent as shown in (13) for /pæ:t/ 'pat', any candidate with a long vowel incurs a violation of LENGTHEN, resulting in the correct prediction of a short vowel.

(13) Input long vowels before voiceless consonants

'pat' /pæ:t/	NOCODA	❀SYM	LENGTHEN	MAX	ID[weight]
a. ❀ pæt	*!				*
b. pæ:t	*!	*	*		
c. ☞ pæ				*	*
d. pæ:		*!	*	*	

Similarly, if an input long vowel were posited before a voiced obstruent in a morphologically complex form (e.g., /dɔ:gi/ 'doggie'), any candidate with a short vowel would be eliminated in favor of a long vowel due to the need to comply with higher ranked LENGTHEN. This is important because it is precisely facts of this sort (the transparent cases) that the brute force account had to treat as accidental. The sympathy account also has the advantage of systematically accounting for the nonoccurrence of a vowel length contrast in words that would not exhibit a consonantal alternation (e.g., *do, doing*). Such words are relevant because the winning candidates would not violate MAX and thus would not be affected by the sympathy constraint. By ranking LENGTHEN above ID[weight], no vowel length contrast could survive in such words, no matter what is assumed underlyingly about length.

This sympathy account is fully consistent with richness of the base by allowing both long and short vowels in all contexts as possible input representations. The constraint hierarchy does all of the work, accounting for the occurrence and nonoccurrence of the various vowel length patterns. Because the account complies with richness of the base, lexicon optimization can also be invoked. Lexicon optimization claims that Child A in this instance with this hierarchy would preferentially internalize an input long vowel for any occurring output long vowel and an input short vowel for any occurring output short vowel. The inputs would be the same given the adult hierarchy. The reason for this choice is that it is precisely these input representations that are the most harmonic relative to the corresponding optimal output representations. To the extent that lexicon optimization is making a correct claim, it must be acknowledged that this principle can only be invoked and lead to this conclusion after the constraint hierarchy has done the work of predicting the output from an unconstrained set of inputs.

The aforementioned sympathy account makes crucial reference to properties of grammar that are presumably universal—namely, the constraints, including the sympathy constraint and the designation of a flower candidate. As such, there should be no cost associated with sympathy. A grammar exhibiting the effect of sympathy cannot be judged as more complicated than a grammar without the same effect. However, if sympathy is universal, why do all grammars not show the effect of sympathy? Such universals must be reconciled against observed individual differences. The characterization of individual differences thus provides for a further test of the theory, which we take up in the following section.

2.6. Individual Differences

Recall that only some of the children from the Weismer et al. (1981) study were observed to exhibit the opacity effect associated with a vowel length distinction before omitted obstruents. Thus, in addition to Child A (and Child B), another child, Child C (age 3;10), also omitted final obstruents but was found to maintain no statistically significant vowel length differences. Vowel durations were highly variable, consistent with the absence of a vowel length contrast and the absence of any contextual conditioning. Such variability is a hallmark of developing systems and is in no way atypical (K. Rice (1996)). The tokens in (14) are from Child C and illustrate the omission of stem-final obstruents and the absence of vowel length differences.

(14) Child C (age 3;10)

a.	dɔ	'dog'	i	'eat'
	dæ	'dad'	peɪ	'plate'
	bɛ	'bed'	tʌ	'truck'
	wɛ	'red'		
b.	dai	'doggie'	iin	'eating'
	dæi	'daddy'	wʌɪn	'looking'

The forms in (14a) show that final obstruents were omitted with no corresponding vowel length differences. The forms in (14b) further reveal that even word-medial obstruents were omitted, suggesting that the omission of final obstruents may have been attributable to a more general prohibition against postvocalic obstruents of any kind (coda or word-medial onset). The facts for this child differed from those of Child A in that vowel length was entirely transparent, being undifferentiated before omitted obstruents, and the omitted obstruents did not alternate (i.e., target stem-final obstruents were realized as null in both base and derived forms).

If we compare our derivational account for Child A with a derivational account of these facts, a couple of possibilities are suggested. On one hand, it might be argued that Child C internalized adult-like underlying representations (similar to Child A and as assumed in much of the acquisition literature) with roughly the same two rules applied in a bleeding order (opposite of Child A). Clearly, the final consonant deletion rule would have had to be formulated more generally in this instance, but it would have applied before the vowel lengthening rule, effectively removing the conditioning environment for lengthening. Under this account, the grammars of the two children would have been the same in terms of underlying representations and nearly so in terms of rules; the difference would have been attributed to differences in rule ordering. Of course, this ordering would have effectively removed any motivation for the vowel lengthening rule because all postvocalic obstruents (even those in medial contexts) would have been deleted prior to the application of the vowel lengthening rule. This suggests that the grammars may instead have differed by the presence versus absence of certain rules. A more concrete account for Child C might have assumed non-adult-like underlying representations that would have been constrained by a morpheme structure condition to exclude postvocalic obstruents. No phonological rule of deletion or lengthening would have been needed. Under this account, the children's grammars would have differed in the substance of their underlying representations and the presence or absence of rules.

Independent of the merits of either derivational alternative, richness of the base within OT demands that these facts be accounted for without constraining the substance of underlying representations. Thus we do not have available within OT the possibility that Child C would have excluded from all of his underlying representations postvocalic obstruents or a vowel length distinction. It is the constraints and constraint rankings that must achieve these effects. Also, because the constraints are presumed to be universal, the grammars for the two children must contain the same constraints. This latter assumption warrants some comment given the difference in the two children's obstruent omission error patterns. These differences can be related to the same constraint if one is recognized as a particular instance of the other. That is, the omission of word-final obstruents achieved by high-ranked NoCoda is possibly just a particular instance of a more general family of constraints disfavoring postvocalic obstruents of any kind. This family of markedness constraints might be formulated as *VC in (15). Such a constraint may find some basis in other phenomena in which material that is not word initial or foot initial is weak or in a nonprivileged context (e.g., Beckman (1997), Zoll (1998)).[11] Whatever the basis or formulation of the constraint, it is not uncommon to observe in early stages of acquisition that words are limited to a single consonant (e.g., Donahue (1986)).

(15) *VC: Avoid postvocalic obstruents.

This entire family of constraints would be undominated in Child C's grammar. The difference for Child A, however, would be that the constraint is exploded with the more specific instance NoCoda, ranked above Max. The other general instances of *VC would be ranked below Max in Child A's grammar. By ranking the constraints in this way for Child A, all postvocalic obstruents would incur a violation of that family of markedness constraints, but some of those violations would be more serious than others. As a result, postvocalic obstruents that are word final would be omitted, but postvocalic obstruents that are word medial would be retained (even though they violate the more general, lower ranked instance of *VC).

Returning to the issue of sympathy in Child C's grammar, the absence of opacity effects along with the absence of long vowels is what might be expected if sympathy played no role. The universalities associated with sympathy, however, can be maintained if a slightly different ranking of all the same constraints were

[11]A possible alternative to *VC is available if OO faithfulness constraints are indeed high ranked in early stages of development, as argued by Hayes (1999) and McCarthy (1999a). That is, the omission of medial consonants would instead follow from a highly ranked OO faithfulness constraint, which would take as its input the base form of the word, which is itself governed by undominated NoCoda. Given the further argument (Hayes (1999), McCarthy (1999a)) that acquisition proceeds by the demotion of markedness or OO faithfulness, or both, the absence of this OO faithfulness effect (i.e., the presence of a consonantal alternation) in Child A's forms would be attributed to the demotion of the OO faithfulness constraint. This alternative does not impinge on the validity of our claims about sympathy.

assumed. This means that the low-ranked faithfulness constraint MAX would still designate some candidate as the object of sympathy and that ❀SYM would incur a violation from any candidate that failed to preserve the vowel length of the flower candidate. In this instance, however, it is apparently more important to comply with some constraint other than ❀SYM, as we see next. The tableau in (16) adopts a minimally different ranking of the constraints and considers a representative target word ending in a voiced obstruent—for example, *dog,* for Child C.

(16) Lower ranked sympathy results in a transparent output

'dog' /dɔg/	*VC	LENGTHEN	❀SYM	MAX	ID[weight]
a. dɔg	*!	*	*		
b. ❀ dɔːg	*!				*
c. ☞ dɔ			*	*	
d. dɔː		*		*	*!

Our account of Child C assumes that the relevant markedness constraints outrank these particular faithfulness constraints. We also assume here that the sympathy constraint dominates the IO faithfulness constraints, although it is acknowledged that sympathy might appear to be even lower ranked in such cases. We see, however, that this assumption has important implications for the characterization of the course of development. As in the case of Child A, we do not need to make any assumptions about an underlying vowel length distinction. Because LENGTHEN dominates ID[weight], any assumptions about input vowel length will give way to markedness, resulting in noncontrastive vowel length and the phonetic occurrence of short vowels in all contexts for Child C. The flower candidate is identified in the same way as for Child A. That is, low-ranked MAX defines the set including candidates (a) and (b) with final consonants. Of those, candidate (b) with the long vowel before a voiced consonant would be more harmonic, serving as the object of sympathy. In the evaluation of candidates, the ranking of *VC above MAX (and the other constraints) results in the elimination of candidates (a) and (b). The remaining two candidates, (c) and (d), each violate either ❀SYM or LENGTHEN, which can be equally ranked if a short vowel is assumed in the input, resulting in a tie. That is, candidate (d) with a long vowel violates LENGTHEN but complies with ❀SYM; candidate (c) with a short vowel violates ❀SYM but complies with LENGTHEN. The choice between the two must then be passed down to a lower ranked constraint. Although both candidates violate MAX, candidate (d) with a long vowel further violates ID[weight] (if a short vowel were assumed for the input representation). The transparent candidate (c) with a short vowel is thus selected as optimal. If for some reason (e.g., richness of the base) an underlying long vowel were alternatively assumed, the desired output with a short vowel would still emerge by simply ranking LENGTHEN over ❀SYM. The important point is that for Child C ❀SYM cannot be ranked above LENGTHEN, and it does not need to be ranked below any of the particular IO faithfulness constraints.

One point of this demonstration has been that the absence of opacity and overt sympathy effects for some children (e.g., Child C) in no way negates the universal character of such constraints. Rather, the individual differences associated with these effects can be attributed to conventional and minimal differences in constraint rankings. Our appeal to sympathy in the case of Child C was not crucial but was instead intended to show that sympathy is not compromised by transparent cases. Whether or not sympathy is adopted, all optimality accounts can achieve the desired results for Child C without any limitations on input representations as long as the relevant markedness constraints outrank the particular faithfulness constraints.

The validity of our accounts of these individual differences can be further evaluated by considering them in the context of a probable course of development that could have led to such differences. The accounts employing a sympathy constraint (whether high ranked for Child A or lower ranked for Child C) must fit within some developmental progression. Similarly, accounts disallowing sympathy (such as the brute force account) must also fit within a developmental progression. It is generally assumed that acquisition proceeds in incremental steps by building on successive grammars (Ingram (1989a)). Thus it is expected that a high degree of continuity should be preserved in successive grammars. In the following section, we consider how well competing accounts deal with the continuity issue.

2.7. Stages of Development

Our characterization of the individual differences associated with these interacting phenomena offers some insight into how opacity effects might emerge and how certain faithfulness constraints might come to outrank markedness constraints in fully developed systems. These insights are moreover suggestive of a likely developmental progression. We argue that an OT employing sympathy preserves a greater degree of continuity across the grammars at different stages and thus accrues a further measure of support.

As noted previously, early stages of development with their many production errors have been characterized by ranking the relevant markedness constraints over particular faithfulness constraints. Fully developed systems with few phonological rules or alternations, on the other hand, tend toward just the reverse ranking. Naturally, we would like to know how these two extremes are related. The assumption has been that this obtains by the incremental demotion of markedness constraints (Tesar and Smolensky (1998); cf. Hale and Reiss (1998)). Yet, how, why, and when do opacity effects emerge in this progression?

An answer is offered if we adopt the further hypothesis that sympathy outranks IO faithfulness at least in the initial state and possibly throughout all stages of development. This hypothesis is consistent with other similar hypotheses about the dominance of OO faithfulness over IO faithfulness in the initial state (Hayes (1999), McCarthy (1999a); cf. Pater (1998)) and various other harmonic (or universal) ranking relations among certain types of constraints (e.g., Kiparsky (1994), Pulleyblank (1997), Smolensky (1995)). A possible rationale for ranking sympathy above IO faithfulness may lie, in part, in sympathy seeming to be a specific instance of faithfulness. That is, the substance of a sympathy constraint tends to mirror the substance of a particular lower ranked IO faithfulness constraint. In this case, both types of constraints preserve vowel length (or weight). As such, sympathy and IO faithfulness might be expected to participate in the special–general relation dictated by the Elsewhere Condition (Kiparsky (1973)). Also, if sympathy were inherently (and permanently) ranked above its counterpart IO faithfulness constraint, the incremental demotion of certain markedness constraints should lead naturally to the emergence of opacity effects. That is, as markedness constraints are demoted, first becoming dominated by sympathy (but not yet by IO faithfulness), the effects of sympathy would become evident. Finally, when the markedness constraints come to be dominated by IO faithfulness (and necessarily by sympathy), the sympathy constraint would be rendered inert and the opacity effects would be eliminated. One reason for this is that there would be no relevant low-ranked faithfulness constraint that could serve as a selector constraint. If there were no selector constraint, there would be no associated flower candidate relevant to the undominated sympathy constraint. Even if there were a flower candidate (identified by a high-ranked faithfulness constraint under an extended view of selectors), that flower candidate would be identical to the input representation, rendering sympathy and IO faithfulness indistinguishable in their effects.

One virtue of imposing the universal ranking of sympathy over its IO faithfulness counterpart is that certain variations in constraint rankings are precluded, in particular, those rankings that have no different empirical consequences. For example, any ranking of IO faithfulness over the relevant sympathy constraint will fail to evidence an effect from that sympathy constraint. We have already seen that sympathy is not always evident (e.g., Child C and adult English); it is thus empirically necessary for the theory to provide for the absence of such effects. Yet how many different rankings do we want to allow to achieve the same effect? We suggest that the single ranking of sympathy over IO faithfulness is sufficient.

Some of the previously mentioned points can be illustrated by a reconsideration of our characterization of individual differences. To the extent that cross-sectional variation corresponds to different stages of development, we should be able to place the grammars of the two children (A and C) along the continuum connecting the extremes of early development and fully developed systems. For example, it might be expected that the chronologically younger Child C represents a relatively early stage of development, with the somewhat older Child A representing a more advanced but intermediate stage. This assumption is supported by other factors as well. That is, aside from age considerations, Child C's productions more poorly match those of the target system in terms of both vowel length and the presence of postvocalic obstruents. Along the same lines, Child A's productions more closely resemble those of the target system, differing primarily in the presence of final obstruents. Child A's realization of a vowel length distinction before omitted final obstruents might be interpreted as a possible alternative phonologization of the underlying voice distinction in those omitted obstruents. Child C shows no evidence of these target distinctions. A close resemblance with the target system in terms of production facts likely corresponds with a close resemblance in grammars. We thus speculate that a grammar similar to that of Child C might change into a grammar similar to that of Child A, which would then change to that of adult English.

This hypothesized developmental progression will have to await verification from case studies that report relevant longitudinal evidence. The claim is nevertheless readily falsifiable by its exclusion of other logically possible developmental sequences. This is not to say that each stage must be directly observed, but rather that the order of the stages cannot be permuted. The general characteristics of this scenario are schematized in (17) in optimality theoretic terms.

(17) Stages of development

Stage 1: Markedness >> Sympathy >> IO Faith
(e.g., Child C, transparent outputs)

Stage 2: Sympathy >> Markedness >> IO Faith
(e.g., Child A, opaque outputs)

Stage 3: Sympathy >> IO Faith >> Markedness
(e.g., archetypical fully developed language, transparent outputs)

Under this scenario, Stage 1 (exemplified by Child C) would be characterized by particular markedness constraints (LENGTHEN and *VC) being ranked above sympathy, which in turn must be ranked above the relevant IO faithfulness constraints. This would result in early outputs being relatively unmarked and transparent. That is, vowel length would be undifferentiated in all contexts, and postvocalic obstruents would be omitted. The important point is that sympathy is ranked as high as possible in the constraint hierarchy consistent with the facts and always above the relevant IO faithfulness constraint. Consistent with richness of the base, the subsequent Stage 2 (exemplified by Child A) would be characterized by the same input representations from the prior stage and would retain the ranking of many of those same markedness constraints over the faithfulness constraints but would demote slightly some of the markedness constraints. More specifically, the emergence of medial obstruents in Child A's grammar would require the demotion of at least part of the *VC constraint. That is, by exploding *VC into its component constraints, the more specific instance of the constraint NOCODA would continue to be ranked above MAX, but the more general instance of the constraint *VC would be demoted below MAX. In addition, the emergence of long vowels before (omitted) voiced obstruents would require LENGTHEN to be demoted below ❀SYM but not yet below IO faithfulness. This new ranking is the only ranking that would ensure the occurrence of long vowels and some opaque outputs. The demotion of LENGTHEN below ❀SYM is the most minimal change in ranking that would allow for the introduction of long vowels in the absence of the conditioning obstruent. Although there is controversy over precisely what facts cause a child to rerank his or her constraints (e.g., Hale and Reiss (1998), Pater (1998), Smolensky (1996a; 1996b)), it may be that the demotion of markedness below sympathy is a necessary precursor to the further demotion of markedness below IO faithfulness. Stage 2 with its opacity effects as documented here may thus constitute a crucial intermediate step toward the ultimate dominance of faithfulness.

The facts of adult English with predictable vowel length and the nonalternating occurrence of postvocalic consonants is indicative of yet another transitional stage between Stages 2 and 3. That is, adult English would entail the further demotion of the markedness constraint NOCODA below the faithfulness constraint MAX. The other markedness constraint LENGTHEN would necessarily continue to dominate ID[weight] in order to ensure the appropriate vowel length in transparent cases. The sympathy constraint would remain undominated, but now it would be rendered inert for lack of a flower candidate different from the input representation. It may be that there is another transitional stage between Stage 2 and adult English where NOCODA would be demoted minimally, becoming unranked relative to MAX. The equal ranking of the two constraints would predict some free variation in the presence or absence of final obstruents. A word such as *cab* might be expected to be realized with and without the final obstruent, but the vowel would be long in either case. This predicted type of variation in the transition between stages is consistent with observed variation in the acquisition of other contrasts for both first- and second-language learners (e.g., Gass (1984), Macken and Barton (1980)). Finally, Stage 3 would represent an archetypical fully developed system in which a considerable number of IO faithfulness constraints (and necessarily sympathy) would outrank the majority of the markedness constraints, yielding transparent and contrastive outputs. The sympathy constraint would continue to be undominated and inert for the reasons just noted.

The optimality theoretic account that admitted sympathy provided for a plausible account of development in which input representations remained unchanged over time, and it was only constraint rankings that changed. The changes in ranking were, moreover, minimal and unidirectional. The ultimate demotion of markedness below faithfulness seemed to come about from the prerequisite demotion of markedness below sympathy. Opacity effects thus arose quite naturally in the course of development. A high degree of continuity was preserved across stages of development. This characterization was consistent with richness of the base and widely held assumptions about the nature of children's underlying representations (relative to the target system) and the dominance of many markedness constraints in early stages of development.

The alternative optimality account that disallowed sympathy (i.e., the brute force account) would be forced to characterize the course of development rather differently. First, the change in grammar from that of Child C to that of Child A and then to adult English would entail anomalous changes in underlying representations as well as in constraint rankings. The earliest stage exemplified by Child C would be characterized much the same under any version of optimality theory. That is, no particular assumption about input vowel length would be necessary, as long as the markedness constraints outranked the particular faithfulness constraints. Also, the nonoccurrence of postvocalic obstruents could be handled as in other accounts—that is, by the dominance of *VC.

The account of the next stage of development (exemplified by Child A) would be comparable only in its treatment of the emergence of word-medial obstruents. The critical difference would arise in the characterization of the emergence of long vowels. The brute force account would require that the constraints be reranked with ID[weight] dominating LENGTHEN and that the input representations be restructured to allow the defective distribution of long and short vowel phonemes. This latter assumption would violate richness of the base and would fail to explain the nonoccurrence of certain patterns—for example, the absence of long vowels before voiceless consonants, the absence of short vowels before voiced consonants, and the absence of a vowel length contrast in words with no postvocalic obstruents. The assumption of an underlying length distinction would also be at odds with the more widely held assumption that children's underlying representations are target appropriate. Nevertheless, for Child A's grammar to change to that of adult English would entail a further change in the constraint rankings and a change in the assumptions about underlying representations, especially if the generalization were to be captured that vowel length is entirely predictable and transparent in English. That generalization would require a return to the original ranking of the relevant constraints—that is, the markedness constraint LENGTHEN must again come to dominate ID[weight]. Such a ranking moreover renders any assumptions about input vowel length distinctions unnecessary. These constraint rerankings would not be unidirectional, violating other widely held assumptions about the nature of change. It is also unclear what fact would motivate a child to demote a faithfulness constraint below a markedness constraint. The standard assumption is that children require positive evidence for the learning or reranking of constraints. The demotion of faithfulness below markedness would instead require the child to take note of the absence of evidence.

Throughout the hypothesized stages in which sympathy was employed, assumptions about the input representations remained unchanged and were consistent with richness of the base. Only the constraint rankings changed over time. The direction of change was consistent with the unidirectional demotion of the markedness constraints motivated on the basis of positive evidence alone, and this led to the dominance of a particular type of faithfulness—namely, the dominance of sympathy. The incremental demotion of the markedness constraints also provided for gradual change and some free variation along the way. All of this is to be preferred over alternative optimality theoretic accounts and scenarios that preserve less continuity over time or that violate other widely held assumptions.

2.8. Representativeness

The case studies considered in this article are admittedly few and are drawn exclusively from a clinical population of children with phonological delays. To be confident about the representativeness of these cases and to extend their implications to normal development would require a broader sample, preferably including longitudinal case studies of normal or disordered development, or both. At least some progress toward this end is provided by other cross-sectional case studies reporting comparable findings for children with phonological delays. For instance, Weismer (1984, 35–41) supplemented the original study with two additional participants, Child D (age 3;10) and Child E (age 4;0), both of whom omitted final obstruents. Vowel length measurements before omitted final obstruents were found to be significantly different in accord with the voicing of the omitted obstruent. These two children thus exhibited the same opacity effect observed for Child A (and Child B). In a similar instrumental investigation, Smit and Bernthal (1983) identified five children with phonological delays (ages 4;6–5;5) who omitted final obstruents. They found that four of the five children exhibited an opacity effect by maintaining statistically significant vowel length differences before those omitted obstruents, preserving the voice contrast from those omitted obstruents. The fifth child (Participant 8) exhibited a transparency effect similar to Child C in that no vowel length distinction was observed before the omitted obstruents.

Similar developmental opacity–transparency effects have been documented for the interaction of vowel lengthening and word-final devoicing. The devoicing process can render lengthening opaque in much the same way that final consonant deletion can. Catts and Jensen (1983), for example, identified nine children with phonological delays (ages 3;10–5;7) who devoiced final obstruents. They found that all but one exhibited an opacity effect by maintaining a vowel length distinction in accord with underlying (but not surface) voicing of final obstruents. The same opacity effect was observed for normal development by Velten (1943). It appears that this particular opacity effect can also persist into the grammars of fully developed languages—for example, German (Port and O'Dell (1984)), Catalan (Dinnsen and Charles-Luce (1984)), and Polish (Slowiaczek and Dinnsen (1985)).

In sum, the combined result from the available case studies that documented the omission of word-final obstruents and happened to include instrumental measures of vowel length was that 8 of the 10 children evidenced an opaque interaction. Similarly, those studies documenting the occurrence of word-final devoicing errors with instrumental measures of preceding vowel length found that 9 of the 10 children exhibited an opaque interaction. It is not entirely clear what to make of the high incidence of opacity effects in these developing systems. Samplings of younger children (with phonological delays) might have yielded just the reverse trend. All that can be concluded is that such opacity effects can and do occur in developing systems. One crucial piece of information that remains missing is true longitudinal evidence about the development of opacity effects. On the basis of the available cross-sectional data alone, we can only speculate about the probable course of development. It is important for future research to document the actual course of development. There is also a striking paucity of evidence from normal development that would bear on the issues raised here. Most studies of normal development have not combined instrumental measures with segmental error analyses as has been done with disordered populations. One reason for this may be that most normally developing children acquiring English establish appropriate vowel length and coda consonants at a relatively early age. At that point, they would have already passed through the stages that would be relevant to the issues raised here. To impose the same experimental task demands on even younger children may make any appropriate study difficult or impossible. One advantage of addressing these issues through the examination of the speech of children with phonological delays is that their somewhat older ages allow them to manage the experimental tasks, often affording a slow-motion view of early development. There is, of course, always the concern of whether there might not be some organic or cognitive basis for the delay. Although this cannot be ruled out, there is little or no qualitative difference (beyond age considerations) between normal and delayed phonological development (e.g., Ingram (1989b), Leonard (1992)). A similar observation has been made about morphosyntactic acquisition by children with specific language impairment such that their grammars resemble those of younger nonimpaired children (e.g., M. Rice, Wexler, and Cleave (1995), Wexler, Schütze, and Rice (1998)).

3. DISCUSSION

At least three other issues are raised by our sympathy account and warrant some further comment here. The first is what we dub the *bouquet problem,* and the second is the role of sympathy for other opacity effects in acquisition. The third issue relates to how OT with sympathy might compare with accounts from derivational theories or other approaches.

3.1. The Bouquet Problem

In our sympathy account mentioned previously, one candidate was designated as the flower candidate in each tableau. However, this has been an oversimplification because there is, in fact, at least one other potential flower candidate in each tableau. The issue arises because there are multiple low-ranked faithfulness constraints that could potentially serve as selector constraints. For example, consider again the tableau in (10) from our sympathy account of the opacity effects for Child A. That tableau is reproduced as (18), with the second flower candidate identified.

(18) Multiple flower candidates with high-ranked sympathy

'cab' /kæb/	NoCoda	❀Sym	Lengthen	Max	ID[weight]
a. kæb	*!	*	*		
b. ❀ kæ:b	*!				*
c. ❀ kæ		*!		*	
d. ☞ kæ:			*	*	*

Candidate (b) with the long vowel and final consonant is the flower candidate that had previously been identified by the selector MAX. Candidate (c) with a short vowel and no coda could also be a potential flower candidate if the other low-ranked faithfulness constraint ID[weight] were permitted to serve as a selector. That is, this other faithfulness constraint could conceivably also serve as a selector, defining a different set of potential flower candidates that includes the perfectly faithful candidate (a) and the codaless candidate (c), both with a short vowel. Of the two, candidate (c) is the more harmonic because it complies with undominated NOCODA and LENGTHEN.

The existence of multiple flower candidates constitutes a virtual bouquet. This is potentially a problem if only one candidate must serve as the effective flower candidate for a given sympathy constraint.[12] Depending, then, on which of these two candidates were to serve as the effective flower candidate, ❀SYM would select as optimal either candidate (c) or (d). The designation of the effective flower candidate from the bouquet is critical here because different conflicting empirical predictions would be made about the optimal output. On empirical grounds, we know that candidate (d) must be selected as the optimal output in correspondence with the flower candidate (b), which in turn would have been identified by the selector MAX. So why was MAX the selector rather than ID[weight]? McCarthy (1999b) bypassed the issue by simply stipulating the selector. Although the choice

[12] The bouquet problem as developed here differs from McCarthy's (1999b) treatment of those cases in which there might be multiple flower candidates. In his discussion, each flower candidate is uniquely related to a different opacity effect, which is governed by a different sympathy constraint. There is no potential for contradictory predictions in McCarthy's (1999b) cases. The difference in our case is that there is only one opacity effect and, thus, only one sympathy constraint. That one sympathy constraint would make contradictory predictions given two competing flower candidates identified by different selectors. For an alternative approach to this issue, see Walker (1999).

between flower candidates or selector constraints could be stipulated, it is deemed more desirable to make the choice on principled grounds. Thus we next entertain a possible solution that avoids any restrictions on selectors or flower candidates.

An independently necessary universal principle is available and could be used for the purpose of selecting the one empirically correct flower candidate—namely, the Elsewhere Condition (Kiparsky (1973)). More specifically, if MAX and ID[weight] were both permitted to serve as potential selector constraints, they would identify the two flower candidates (b) and (c), as just described. Moreover, if long vowels were assumed to be represented by a monovalent feature [long] (or by two moras), the representation of flower candidate (b) with a long vowel and final consonant would be seen to properly include the representation of the other flower candidate (c) with a short vowel and no final consonant. Thus, when there are multiple flower candidates, and their representations are in this subset relation, the more specific (properly including) representation is chosen as the effective flower candidate. The Elsewhere Condition would correctly select flower candidate (b) as the effective flower candidate. In addition, because the Elsewhere Condition is a universal principle, the choice in every such case would be uniform and nonstipulative.

The bouquet problem can also arise in the transparent cases in which sympathy would seem to play no overt role. This is illustrated by reconsidering the case of Child C, especially tableau (16) (reproduced here as (19)). Thus, in addition to the flower candidate (b) that was selected by MAX, candidate (c) is another flower candidate selected by ID[weight]. These two flower candidates are in a subset relation, as illustrated in the prior discussion, and would result in candidate (b) being selected as the effective flower candidate, as was assumed in our original discussion in section 2.6. We saw there that candidate (c) was correctly predicted as optimal. Though high-ranked ❀SYM and LENGTHEN could not resolve the tie between candidates (c) and (d), the lower ranked faithfulness constraint was decisive in selecting candidate (c) as optimal.

(19) Multiple flower candidates in transparent cases

'dog' /dɔg/	*VC	LENGTHEN	❀SYM	MAX	ID[weight]
a. dɔg	*!	*	*		
b. ❀ dɔːg	*!				*
c. ☞ ❀ dɔ			*	*	
d. dɔː		*		*	*!

A preliminary solution to the bouquet problem has been advanced here through appeal to a universal principle that evaluates the inclusion relation between representations of competing candidates. No stipulations were needed about the selector or the flower candidate. It remains to be determined whether the bouquet problem can be resolved in the same way in all cases of sympathy. Especially relevant would be cases in which the representations of competing flower candidates are not in a subset relation. If such cases exist, it may be necessary to eliminate some potential flower candidates by further restricting or stipulating what is a possible selector constraint. Such an approach might follow along the lines originally suggested by McCarthy (1999b) but must be acknowledged to increase the power of the theory, possibly beyond prudent limits.

3.2. Other Cases of Opacity

In derivational terms, the opacity effect that has been considered here followed from a counterbleeding interaction. A number of other counterbleeding interactions with similar opacity effects should be possible in developing systems. It might be expected that all such cases would be equally amenable to and supportive of sympathy, but at least some of these cases appear to pose a different set of theoretical challenges. For example, many acquisition studies have documented the phenomenon of consonant harmony in which a consonant takes on the place or manner of a nonadjacent consonant elsewhere in the word (e.g., Vihman (1978)). This phenomenon can interact with other common and independent error patterns, which might either delete a final consonant or replace it with a glottal stop. These other error patterns can render assimilation opaque by obscuring the conditioning environment for assimilation. In such cases, a word such as *top* would be realized as [pa] (but not *[ta]). In derivational terms, the initial coronal consonant would take on the labial place of the final consonant by some rule of assimilation. An additional rule of final consonant deletion would be ordered after assimilation in a counterbleeding order, deleting the consonant that triggered assimilation.

Although a number of optimality theoretic accounts of consonant harmony have been advanced (e.g., Dinnsen and Barlow (1998b), Goad (1997), Pater (1997)), none has attempted to account for this opaque interaction between harmony and final consonant omission. If sympathy were to be employed to handle this interaction, the constraints in (20) might be entertained.

(20) Preliminary constraints and ranking

a. Markedness constraints

AGREE:	Avoid different consonantal place gestures within a word.
*LABIAL:	Avoid labial consonants.
*CORONAL:	Avoid coronal consonants.
NOCODA:	Avoid coda consonants.

b. Faithfulness constraints

ID[labial]:	Corresponding segments must be identical in terms of the feature [labial].
ID[coronal]:	Corresponding segments must be identical in terms of the feature [coronal].
MAX:	Every input segment has a corresponding output segment (no deletion).

c. Sympathy constraint
❀SYM: Corresponding segments (of the flower candidate and an output candidate) must be identical in terms of place features.
Ranking: AGREE, NOCODA, ❀SYM >> ID[lab] >> ID[cor], *LABIAL >> MAX, *CORONAL

Something along the lines of a highly ranked markedness constraint, AGREE, would account for place harmony by favoring output candidates with one (multiply linked) place feature or alternatively two identical consonantal place features within a word. The individual markedness constraints *LABIAL and *CORONAL each assess violations for specific place features. By ranking *LABIAL above *CORONAL, an explanation is offered for why in developing systems velars are often replaced by coronals rather than by labials in nonassimilatory contexts. The familiar markedness constraint NOCODA militates against final consonants. The faithfulness constraints are antagonistic to these markedness constraints. Identity in specific place features for corresponding segments is demanded by the individual constraints ID[labial] and ID[coronal]. The ranking of ID[labial] over ID[coronal] is intended to account for the tendency for labials to serve as triggers of assimilation and the vulnerability of coronals as targets of assimilation. Although MAX militates against deletion, its ranking below NOCODA tolerates final consonant omission. The sympathy constraint ❀SYM demands identity of place features in corresponding segments of the flower candidate and an output candidate.

For a sympathy account to work in such a case, it would be necessary for the most harmonic failed candidate (i.e., the flower candidate selected by MAX) to be [pap] for input /tap/. The sympathy constraint would need to be highly ranked to preserve the correspondence in place features between the initial consonant of the flower candidate and the optimal output candidate [pa]. The problem is that [pap] cannot be the flower candidate for at least two reasons, especially given conventional assumptions about IDENT[feature] constraints (e.g., McCarthy and Prince (1995), Pater (1999)) and universal hierarchies relating to place features (e.g., Kiparsky (1994)).

(21) Failure of sympathy

'top' /tap/	AGREE	NOCODA	❀SYM	ID[lab]	ID[cor]	*LAB	*COR	MAX
a. tap	*!	*	*			*	*	
b. pap		*!	**	*	*	**		
c. ❀ tat		*!		*	*		**	
d. ☞ ta							*	*
e. → pa			*!	*	*	*		*

To see this for input /tap/, consider the tableau in (21) and the two most likely competing flower candidates (b) [pap] and (c) [tat], both of which comply with the low-ranked selector MAX and the highly ranked markedness constraint AGREE. It is this markedness constraint that eliminates the faithful candidate (a) [tap] from the flower candidate competition. The other two competing candidates each incur a single ID[labial] violation, albeit for different reasons. The initial consonant of [pap] includes a labial place feature that is not in the input, and the final consonant of [tat] fails to include the input labial feature. Both candidates similarly incur a single violation of lower ranked ID[coronal]. Thus they are equally unfaithful in terms of place. The choice must then be passed down to the individual markedness constraints. If we adopt the universal ranking of *LABIAL over *CORONAL, more serious violations will be assessed to labials than to coronals. As a result, [pap] is judged less harmonic than [tat], incorrectly selecting [tat] as the flower candidate. The sympathy constraint would then preserve the place of the initial consonant of the flower candidate to yield the erroneous output (d) [ta],[13] when it is in fact (e) [pa] that should win (as indicated by the arrow). It seems, then, that sympathy is incapable of accounting for this particular opacity effect.

If, however, a different set of assumptions were adopted about feature faithfulness, namely MAX[feature] constraints (Lombardi (1998)), the desired result could be achieved in this case without any appeal to sympathy.[14] This alternate interpretation requires only that a particular feature be parsed, but not necessarily in the same corresponding segment. Assume for the moment that the ID[feature] constraints in (20) were replaced by the MAX[feature] constraints in (22). The tableau in (23) reconsiders the evaluation of candidates under this alternative without any appeal to sympathy.

(22) Alternative feature faithfulness constraints
MAX[labial]: Preserve (or parse) the input feature [labial].
MAX[coronal]: Preserve (or parse) the input feature [coronal].

(23) Opaque consonant harmony without sympathy

'top' /tap/	AGREE	NOCODA	MAX[lab]	MAX[cor]	*LAB	*COR	MAX
a. tap	*!	*			*	*	
b. pap		*!		*	**		
c. tat		*!	*			**	
d. ta			*!			*	*
e. ☞ pa				*	*		*

[13]The same incorrect prediction would be made if ID[labial] were undominated and [tap] were chosen as the flower candidate.

[14]We are grateful to two anonymous reviewers for pointing out to us the availability of such an alternative.

Under this view, then, candidates (b) [pap] and (e) [pa] in (23) would not incur MAX[labial] violations because the [labial] place feature of the final consonant is being parsed (albeit in the initial consonant). These same candidates would incur violations of MAX[coronal] for their failure to parse the [coronal] feature of the initial consonant, but the lower ranking of that constraint results in less serious violations. Candidates (c) [tat] and (d) [ta] would incur the more serious violations of MAX[labial] for their failure to parse the [labial] feature of the final input consonant. Given high-ranked NOCODA and the importance of parsing [labial] place features, the winning candidate for input /tap/ would correctly be predicted to be (e) [pa]. It is unclear whether a unified interpretation of feature faithfulness is available to account for the full range of phenomena interacting with other cases of consonant harmony (cf. Dinnsen (1998), Dinnsen, Barlow, and Morrisette (1997)). At the very least, this suggests that (counterbleeding) opacity effects that involve assimilation may be different from other counterbleeding cases and warrant further consideration for their relevance to sympathy. Even if MAX[feature] constraints prove necessary, they are not helpful in the opaque lengthening cases because no feature of the omitted obstruent is being preserved in the vowel, and the vowel lengthens whether the obstruent is omitted or not (cf. lengthening before word-medial obstruents).

It will also be important to examine other opacity effects in acquisition that arise from other types of interactions, including, most notably, counterfeeding relations. Chain shifts (e.g., /θ/ is replaced by [f], but /s/ is replaced by [θ]) constitute one such relevant interaction common in developing systems. Interestingly, however, optimality theoretic accounts of these phenomena are available that do not appeal to sympathy (Dinnsen and Barlow (1998a)), calling into question the necessity of sympathy in at least certain cases of opacity. Although opacity effects associated with chain shifts or consonant harmony, or both, ultimately may not support sympathy, it also cannot be said that they refute sympathy. The identification of other types of counterbleeding and counterfeeding interactions from acquisition should allow for a fuller assessment of the role of sympathy.

3.3. Comparison With Other Accounts and Frameworks

The intent of this article has not been to compare OT with derivational theories, but rather to evaluate competing proposals and their associated accounts of acquisition within the particular framework of OT. Nevertheless, there may be some value in comparing different theoretical frameworks in terms of their claims and assumptions about acquisition. As noted earlier, derivational theories would have little difficulty accounting for the individual differences and opacity–transparency effects associated with Child A and Child C. The more revealing assessment comes from a consideration of derivational claims about the developmental progression. Given that derivational theories allow underlying representations to be constrained in various ways, one possible account of the individual differences would be to exclude long vowels from underlying representations, assuming short vowels throughout the stages of development, and to allow for postvocalic obstruents in accord with the target system. The difference in the grammars of the two children would reside in the order of the rules as well as in the formulation of one of the rules (i.e., the deletion rule). That is, for Child A, the rules would apply in a counterbleeding order with the deletion rule restricted to word-final obstruents. For Child C, the rules would apply in a bleeding order with the deletion rule formulated more generally to include postvocalic obstruents (even those in word-medial contexts).

Although this account of the individual differences would allow some degree of continuity to be preserved in the grammars of these two children, it does so at the cost of introducing other limitations. Specifically, both rules in Child C's grammar would lack such standard motivation as a consonantal alternation and an observable vowel length distinction. It is also unclear from a developmental perspective why the rules would have reordered to introduce the allomorphy observed in Child A's system. This assumes, of course, that Child C's grammar reflects the precursor to Child A's grammar. If the course of development were instead assumed to be reversed, the reordering would have leveled the paradigm, motivating the restructuring of underlying representations and the loss of both rules. The challenge would then be to explain the required changes in underlying representations and the reintroduction of the lengthening rule that would be associated with the transition from Child C's grammar to that of adult English. Any alternative derivational account would preserve less continuity and would violate the widely held assumption that children's underlying representations are adult-like. For example, returning to the more likely scenario that Child C represents the precursor to Child A's grammar, Child C might have been assumed to further restrict his underlying representations by excluding postvocalic obstruents from the underlying representations, obviating any need for either a lengthening rule or a deletion rule. By this derivational account then, the grammars of the two children would have differed in the substance of their underlying representations and the presence or absence of rules. The developmental progression would entail the claim that the underlying representations changed at the same point that two new rules were added to the grammar (to yield Child A's grammar from that of Child C). The further development into adult English would require the loss of the deletion rule. Why rules would be added or lost from grammars has never received a fully satisfying account. A child's addition of the lengthening rule and his subsequent loss of the deletion rule might be seen as an attempt to match the target system, but why then would that same child have ever added the deletion rule when no such rule is motivated for the target system?

It is concerns of this sort that made certain aspects of natural phonology seem attractive (Donegan and Stampe (1979)). That is, lengthening and deletion might have been considered innate natural processes that had to be suppressed (or lost) by the child in the course of development. Although Child C did not exhibit the lengthening effect associated with one of the processes, deletion might have been ordered before lengthening in a bleeding relation. The processes could then have reordered to yield Child A's grammar, which in turn would develop into adult English by the suppression of the deletion process. One problem with this account is that lengthening would likely be classified as a fortition process because it enhances the voice contrast from neighboring sounds, and as such would be prevented from applying after a lenition process such as deletion (Donegan and Stampe (1979, 153–156)). The alternative might be to consider lengthening a learned rule (rather than an innate process), which Child C simply had not yet learned. However, given the near universal character of lengthening, this alternative also seems unattractive.

Finally, another alternative that is available both within derivational theories and OT is to relegate the obstruent omission errors (and especially the opacity effects) to performance factors such as motor immaturity (as was argued by Hale and Reiss (1998) for many other developmental phenomena). The claim in such a case would be that the children's grammars are largely intact and adult-like. Errors arise not in the output of the grammar but rather as the output of the "body." We do not doubt that performance factors play a role in both developing and fully developed systems, but it seems unlikely that the occurrence of these same phenomena in fully developed languages would be attributed to performance limitations. Vowel lengthening and coda restrictions are each independently occurring and prevalent phenomena in developing and fully developed languages. To dismiss the developmental evidence ignores a potentially valuable source of insight into language and does not eliminate the problem for the characterization of fully developed languages.

OT addresses these issues rather differently. Although there are no rules (and thus no issue of rule loss, rule addition, rule reformulation, or rule reordering), constraints and constraint rankings do achieve many of the same effects as rules. However, because the constraints are universal, nothing is lost or added. The presence or absence of some effect is attributed to the ranking of constraints. Default (or initial state) rankings, which are hypothesized to find markedness and OO faithfulness constraints to be highly ranked, account for children's many production errors (relative to the target system) in the early stages of development. If a child is to succeed at eliminating these error patterns, positive evidence will be required to demote the markedness or OO faithfulness constraints, or both. In addition, children's underlying representations are prevented from changing over time given richness of the base and the universal character of underlying representations. Thus, although the widely held assumption of adult-like underlying representations need not be violated, nothing crucial depends on it either, at least in the early stages of development in which many markedness constraints dominate the antagonistic faithfulness constraints. Again, it is the constraint hierarchy that must be able to achieve the effect of limiting underlying representations (lexicon optimization). A high degree of continuity is thus preserved in grammars over time, with only the constraint rankings being permitted to change. The range of variation is further restricted by those cases in which universal rankings are involved, such as that proposed here for the dominance of sympathy over its corresponding IO faithfulness constraint. The requirements of OT thus provide for a high degree of continuity to be preserved in children's grammars as they develop over time.

A comparison of these frameworks finds that derivational theories allow many properties of grammar to vary and change over time. The presence or absence of rules, the formulation of rules, the ordering of rules, and the substance of underlying representations are among the derivational devices that are available to change. In contrast, OT allows grammars to change only in terms of the constraint rankings. It will be difficult to fully appreciate the empirical consequences of these differences until many other issues are resolved, including the discovery of the substantive properties of the universal constraints and the controversy over what can serve as a selector constraint. The frameworks appear to differ in their predictions about the course of development. For example, although some derivational theories predict that Child C could have changed his grammar to that of Child A and then to that of adult English, all other logically possible developmental sequences are also predicted. On the other hand, if (as prescribed by OT) acquisition proceeds by the demotion of markedness constraints, then only certain steps in the developmental progression should be possible, offering an explanation for why Child A's grammar would be intermediate to Child C and adult English. Although this comparison would seem to favor OT as an account of acquisition, it must be recognized that there are still many other fronts on which these theories can and should be tested. What is clear is that OT could not even be in the running, unless sympathy (or something like it) were available to account for Child A and the other comparable opacity effects observed in the cited studies.

4. CONCLUSION

We hope we have shown that the developing systems of young children evidence opacity effects that are similar to those in fully developed languages and that an appeal to sympathy offers an insightful account of at least some of those effects. Acquisition phenomena generally, and the particular interaction of vowel lengthening and final consonant omission, have proven especially relevant to a test of OT, and more specifically to sympathy. One reason for this is that early stages of acquisition present with the crucial conditions that should give rise to opacity effects and presumably require sympathy—namely, a number of low-ranked faithfulness constraints. It is these low-ranked faithfulness constraints that are claimed to identify a flower candidate, making available properties different from the input representation that could be preserved by the higher ranked sympathy constraint. As markedness constraints are demoted in the course of development, opacity effects of the sort documented here should emerge.

The value of sympathy is further supported by its ability to maintain a high degree of continuity in the grammars across stages of development. The constraints and input representations remain unchanged over time with only the ranking of constraints changing. Although the triggers for constraint reranking are still unclear, sympathy does offer some new insight into how and when opacity effects might emerge and how faithfulness might ultimately come to outrank markedness constraints in fully developed systems. That is, the initial demotion of markedness reveals the dominance of sympathy with its attendant opacity effects and only later the dominance of IO faithfulness. The dominance of sympathy over IO faithfulness appears to reflect a harmonic ranking and would seem to provide for a transitional compromise of sorts between the two extremes of undominated markedness and undominated faithfulness. The compromise is complying with markedness and preserving something of a representation that cannot otherwise survive.

In the course of our discussion, many issues have been raised that will require further consideration and empirical validation. Perhaps most central is the need for detailed longitudinal case studies that document the course of development in which interacting phenomena are involved. As such studies become available, it should be possible to determine whether opacity effects of the sort reported here do naturally emerge as intermediate stages of development. Another issue that is central to all acquisition research but that has continued to elude a satisfactory resolution in OT (including our sympathy account and, for that matter, in any theoretical framework) is the comprehension–production dilemma. Whatever the ultimate solution turns out to be, it will remain necessary to provide for the opacity effects evident in the production domain. Sympathy at least provides for an optimality theoretic account of certain opacity effects and their development. There is certainly nothing inherent to sympathy that denies the existence of the comprehension–production dilemma or that should preclude its resolution. At the very least, sympathy as a theoretical construct appears to offer some new insights for both theory and acquisition.

ACKNOWLEDGMENTS

This work was supported by Grants DC01694 and DC00012 from the National Institutes of Health. A preliminary version of this article was presented at the Third Mid-Continental Workshop on Phonology held at Indiana University in October 1997.

The article has undergone a number of significant revisions in response to helpful comments from colleagues. We are especially grateful to four anonymous reviewers, Jessica Barlow, Stuart Davis, Judith Gierut, Junko Itô, Gregory Iverson, Linda Lombardi, John McCarthy, Michele Morrisette, and Paul Smolensky.

REFERENCES

Beckman, J. N. (1997) "Positional Faithfulness, Positional Neutralisation and Shona Vowel Harmony," *Phonology* 14, 1–46.

Benua, L. (1995) "Identity Effects in Morphological Truncation," in J. N. Beckman, L. W. Dickey, and S. Urbanczyk, eds., *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory,* GLSA, University of Massachusetts, Amherst.

Catts, H. W. and P. J. Jensen (1983) "Speech Timing of Phonologically Disordered Children: Voicing Contrast of Initial and Final Stop Consonants," *Journal of Speech and Hearing Research* 26, 501–510.

Chen, M. (1970) "Vowel Length Variation as a Function of the Voicing of Consonant Environment," *Phonetica* 22, 129–159.

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English,* Harper & Row, New York.

Clements, G. N. (1986) "Compensatory Lengthening and Consonant Gemination in LuGanda," in L. Wetzels and E. Sezer, eds., *Studies in Compensatory Lengthening,* Foris, Dordrecht, The Netherlands.

Davis, S. (1997) "A Sympathetic Account of Nasal Substitution in Ponapean Reduplication," in R. Walker, M. Katayama, and D. Karvonen, eds., *Phonology at Santa Cruz* 5, Linguistics Research Center, Santa Cruz, California.

de Lacy, P. (1998) "Sympathetic Stress," ms. [Rutgers Optimality Archive 294–0199]

Dinnsen, D. A. (1998) "On the Organization and Specification of Manner Features," *Journal of Linguistics* 34, 1–25.

Dinnsen, D. A. and J. A. Barlow (1998a) "On the Characterization of a Chain Shift in Normal and Delayed Phonological Acquisition," *Journal of Child Language* 25, 61–94.

Dinnsen, D. A. and J. A. Barlow (1998b) "Root and Manner Feature Faithfulness in Acquisition," in A. Greenhill, M. Hughes, H. Littlefield, and H. Walsh, eds., *Proceedings of the 20th Annual Boston University Conference on Language Development,* Cascadilla Press, Somerville, Massachusetts.

Dinnsen, D. A., J. A. Barlow, and M. L. Morrisette (1997) "Long-Distance Place Assimilation With an Interacting Error Pattern in Phonological Acquisition," *Clinical Linguistics & Phonetics* 11, 319–338.

Dinnsen, D. A. and J. Charles-Luce (1984) "Phonological Neutralization, Phonetic Implementation and Individual Differences," *Journal of Phonetics* 12, 49–60.

Donahue, M. (1986) "Phonological Constraints on the Emergence of Two-Word Utterances," *Journal of Child Language* 13, 209–218.

Donegan, P. J. and D. Stampe (1979) "The Study of Natural Phonology," in D. A. Dinnsen, ed., *Current Approaches to Phonological Theory,* Indiana University Press, Bloomington.

Gass, S. (1984) "Development of Speech Perception and Speech Production Abilities in Adult Second Language Learners," *Applied Psycholinguistics* 5, 51–74.

Gnanadesikan, A. E. (1996) "Child Phonology in Optimality Theory: Ranking Markedness and Faithfulness Constraints," in A. Stringfellow, D. Cahana-Amitay, E. Hughes, and A. Zukowski, eds., *Proceedings of the 20th Annual Boston University Conference on Language Development,* Cascadilla Press, Somerville, Massachusetts.

Goad, H. (1997) "Consonant Harmony in Child Language: An Optimality Theoretic Account," in S. J. Hannahs and M. Young-Scholten, eds., *Focus on Phonological Acquisition,* John Benjamins, Amsterdam, The Netherlands.

Goldrick, M. and P. Smolensky (1999) "Opacity and Turbid Representations in Optimality Theory," paper presented at the 35th meeting of the Chicago Linguistic Society, University of Chicago.

Hale, M. and C. Reiss (1998) "Formal and Empirical Arguments Concerning Phonological Acquisition," *Linguistic Inquiry* 29, 656–683.

Hayes, B. (1999) "Phonological Acquisition in Optimality Theory: The Early Stages," ms. [Rutgers Optimality Archive 327–0699]

Hubbard, K. (1995) "Prenasalized Consonants and Syllable Timing: Evidence from Runyambo and Luganda," *Phonology* 12, 235–256.

Ingram, D. (1989a) *First Language Acquisition: Method, Description and Explanation,* Cambridge University Press, New York.

Ingram, D. (1989b) *Phonological Disability in Children,* Cole & Whurr, London.

Itô, J. and A. Mester (1997) "Sympathy Theory and German Truncations," in V. Miglio and B. Morén, eds., *University of Maryland Working Papers in Linguistics 5: Proceedings of the Hopkins Optimality Workshop/Maryland Mayfest,* University of Maryland, College Park.

Itô, J. and A. Mester (1999) "On the Sources of Opacity in OT: Coda Processes in German," ms. [Rutgers Optimality Archive 347–0999]

Itô, J., A. Mester, and J. Padgett (1995) "Licensing and Underspecification in Optimality Theory," *Linguistic Inquiry* 7, 341–351.

Keating, P. A. (1979) *A Phonetic Study of a Voicing Contrast in Polish,* Doctoral dissertation, Brown University, Providence, Rhode Island.

Keating, P. A. (1985) "Universal Phonetics and the Organization of Grammars," in V. A. Fromkin, ed., *Phonetic Linguistics,* Academic Press, New York.

Kiparsky, P. (1973) " 'Elsewhere' in Phonology," in S. R. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle,* Holt, Rinehart and Winston, New York.

Kiparsky, P. (1994) "Remarks on Markedness," ms., Stanford University.

Kirchner, R. (1996) "Synchronic Chain Shifts in Optimality Theory," *Linguistic Inquiry* 27, 341–351.

Koontz-Garboden, A. (2000) "Opacity in Batticaloa Creole Portuguese Stress Assignment: Motivation for Candidate-to-Candidate Faithfulness," ms. [Rutgers Optimality Archive 395–06100]

Lee, M. (1999) "A Case of Sympathy in Javanese Affixation," in K. Baertsch and D. A. Dinnsen, eds., *Optimal Green Ideas in Phonology: Indiana University Working Papers in Linguistics* 1, Indiana University Linguistics Club, Bloomington.

Leonard, L. (1992) "Models of Phonological Development and Children With Disorders," in C. A. Ferguson, L. Menn, and C. Stoel-Gammon, eds., *Phonological Development: Models, Research, Implications,* York Press, Timonium, Maryland.

Locke, J. (1983) *Phonological Acquisition and Change,* Academic Press, New York.

Lombardi, L. (1998) "Evidence for MaxFeature Constraints From Japanese," ms. [Rutgers Optimality Archive 247–0298]

Łubowicz, A. (1998) "Derived Environment Effects in OT," ms. [Rutgers Optimality Archive 239–0198]

Macken, M. A. and D. Barton (1980) "A Longitudinal Study of the Working Contrast in American English Word-Initial Stops, as Measured by Voice-Onset Time," *Journal of Child Language* 7, 41–74.

McCarthy, J. (1999a) "Morpheme Structure Constraints and Paradigm Occultation," in C. Gruber, D. Higgins, K. Olson, and T. Wysocki, eds., *CLS 34,* Chicago Linguistic Society, Chicago.

McCarthy, J. (1999b) "Sympathy and Phonological Opacity," *Phonology* 16, 331–399.

McCarthy, J. (1999c) "Sympathy, Cumulativity, and the Duke-of-York Gambit," in K. Baertsch and D. A. Dinnsen, eds., *Optimal Green Ideas in Phonology: Indiana University Working Papers in Linguistics* 1, Indiana University Linguistics Club, Bloomington.

McCarthy, J. and A. Prince (1995) "Faithfulness and Reduplicative Identity," in J. N. Beckman, L. W. Dickey, and S. Urbanczyk, eds., *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory,* GLSA, Amherst, Massachusetts.

McGarrity, L. W. (1999) "A Sympathy Account of Multiple Opacity in Wintu," in K. Baertsch and D. A. Dinnsen, eds., *Optimal Green Ideas in Phonology: Indiana University Working Papers in Linguistics* 1, Indiana University Linguistics Club, Bloomington.

Pater, J. (1997) "Minimal Violation and Phonological Development," *Language Acquisition* 6, 201–253.

Pater, J. (1998) “From Phonological Typology to the Development of Receptive and Productive Phonological Competence: Applications of Minimal Violation,” ms. [Rutgers Optimality Archive 296–0299]

Pater, J. (1999) “Austronesian Nasal Substitution and Other NC Effects,” in R. Kager, H. van der Hulst, and W. Zonneveld, eds., *The Prosody Morphology Interface,* Cambridge University Press, Cambridge, Massachusetts.

Port, R., S. Al-Ani, and S. Maeda (1981) “Temporal Compensation and Universal Phonetics,” *Phonetica* 37, 235–252.

Port, R. and M. O’Dell (1984) “Neutralization of Syllable Final Voicing in German,” *Research in Phonetics* 4, 93–134.

Prince, A. and P. Smolensky (1993) “Optimality Theory: Constraint Interaction in Generative Grammar,” Rutgers Center for Cognitive Science (RCCS) Tech. Rep. No. 2, RCCS, Rutgers University, New Brunswick, New Jersey.

Pulleyblank, D. (1997) “Optimality Theory and Features,” in D. Archangeli and D. T. Langendoen, eds., *Optimality Theory: An Overview,* Blackwell, Cambridge, Massachusetts.

Pyle, C. (1971) *The Treatment of Length in Generative Phonology,* Doctoral dissertation, University of Illinois, Urbana–Champaign.

Raphael, L. J., M. F. Dorman, and D. Geffner (1980) “Voicing-Conditioned Durational Differences in Vowels and Consonants in the Speech of Three- and Four-Year-Old Children,” *Journal of Phonetics* 8, 335–342.

Rice, K. (1996) “Aspects of Variability in Child Language Acquisition,” in B. Bernhardt, J. Gilbert, and D. Ingram, eds., *Proceedings of the UBC International Conference on Phonological Acquisition*, Cascadilla Press, Somerville, Massachusetts.

Rice, M., K. Wexler, and P. Cleave (1995) “Specific Language Impairment as a Period of Extended Optional Infinitive,” *Journal of Speech and Hearing Research* 38, 850–863.

Slowiaczek, L. M. and D. A. Dinnsen (1985) “On the Neutralizing Status of Polish Word-Final Devoicing,” *Journal of Phonetics* 13, 325–341.

Smit, A. B. (1993) “Phonologic Error Distributions in the Iowa–Nebraska Articulation Norms Project: Consonant Singletons,” *Journal of Speech and Hearing Research* 36, 533–547.

Smit, A. B. and J. E. Bernthal (1983) “Voicing Contrasts and Their Phonological Implications in the Speech of Articulation-Disordered Children,” *Journal of Speech and Hearing Research* 26, 486–500.

Smolensky, P. (1995) “On the Structure of the Constraint Component Con of UG,” ms. [Rutgers Optimality Archive 86–0000]

Smolensky, P. (1996a) “The Initial State and ‘Richness of the Base’ in Optimality Theory,” Tech. Rep. No. JHU–CogSci–96–4, Department of Cognitive Science, Johns Hopkins University, Baltimore.

Smolensky, P. (1996b) “On the Comprehension/Production Dilemma in Child Language,” *Linguistic Inquiry* 27, 720–731.

Sprouse, R. (1998) “Enriched Inputs Sets: A New Approach to Opacity in OT,” paper presented at the 1998 Annual Meeting of the Linguistics Society of America, New York.

Stoel-Gammon, C. and C. Dunn (1985) *Normal and Disordered Phonology in Children,* University Park Press, Baltimore.

Tesar, B. and P. Smolensky (1998) “Learnability in Optimality Theory,” *Linguistic Inquiry* 29, 229–268.

Velten, H. V. (1943) “The Growth of Phonemic and Lexical Patterns in Infant Language,” *Language* 19, 281–292.

Vihman, M. (1978) “Consonant Harmony: Its Scope and Function in Child Language,” in J. Greenberg, ed., *Universals of Human Language* 2, Stanford University Press, Stanford, California.

Walker, R. (1999) “Reinterpreting transparency in nasal harmony,” ms. [Rutgers Optimality Archive 306–0399]

Weismer, G. (1984) “Acoustic Analysis Strategies for the Refinement of Phonological Analysis,” in M. Elbert, D. A. Dinnsen, and G. Weismer, eds., *Phonological Theory and the Misarticulating Child,* American Speech–Language Hearing Association Monographs Number 22, Rockville, Maryland.

Weismer, G., D. A. Dinnsen, and M. Elbert (1981) “A Study of the Voicing Distinction Associated With Omitted Word-Final Stops,” *Journal of Speech and Hearing Disorders* 46, 320–328.

Wexler, K., C. T. Schütze, and M. Rice (1998) “Subject Case in Children With SLI and Unaffected Controls: Evidence for the Agr/Tns Omission Model,” *Language Acquisition* 7, 317–344.

Zoll, C. (1998) “Positional Asymmetries and Licensing,” ms. [Rutgers Optimality Archive 282–0998]