Sympathy, Cumulativity, and the Duke-of-York Gambit[*]

John J. McCarthy

University of Massachusetts, Amherst

1 Introduction

Oh, the grand old Duke of York,
He had ten thousand men;
He marched them up to the top of the hill,
And he marched them down again.

An English Nursery-Rhyme

Serial derivations have been a central idea in the theory of generative phonology throughout its history, but scant attention has been paid to a key question: is any serial derivation possible in human languages? More precisely, can any independently licit rule co-exist with any other licit rule, and can the rules apply in any order? The rule co-existence question has, to my knowledge, never been raised in the literature. The rule ordering question was investigated intensively during the early 1970's (see Iverson (1995) for a review), but often just a pair of rules was studied in isolation from the broader derivational context. Since about 1975, research in generative phonology has mostly dealt with the form of rules and the nature of representations — subjects which are interesting in themselves but do not help advance the theory of derivations.

A rare contribution to this neglected topic is Pullum's (1976) study of the "Duke-of-York gambit". Duke-of-York (DY) derivations have the general form A→B→A, where underlying A passes through a B stage before returning to surface A again. For example, in some analyses of *r*-dropping and *r*-intrusion in various English dialects, final *r* is first deleted and then re-inserted before a vowel: *Homer is* → *Hom[ə] is* → *Homer is* (cf. *Hom[ə] saw*). Pullum addresses this case and others like it, asking whether DY derivations are required by the facts and how they might be ruled out generally. (In his view, they are required and should not be ruled out. More on this below.)

Optimality Theory (Prince and Smolensky 1993) offers a novel perspective on process co-existence and interaction. It is to be expected, therefore, that OT can yield new insights into DY derivations and, by extension, into the questions posed at the outset of this article. I therefore propose to re-visit the topic of DY derivations within the context of OT.

There are two main types of DY derivations, and they turn out to have very different implications for linguistic theory. In the first type, which I call *vacuous*, the intermediate stage of the A→B→A derivation has a somewhat artifactual status, as in the following hypothetical example:

(1) Vacuous DY Derivation

	Underlying	/CAD/	cf.	/ZAD/	/CBW/
⇨	A→B/__D	CBD		ZBD	—
⇨	B→A/C__	CAD		—	CAW

The last two columns show that both rules are independently motivated; the focus is on the column headed by /CAD/. The DY derivation /CAD/ → *CBD* → *CAD* is vacuous because nothing else depends on the intermediate stage CBD. The theory-internal assumptions of strict serialism, rather than some empirical argument, motivate this intermediate stage. It is easily dispensed with in OT, as I will show below.[1]

In *feeding* DY derivations, the intermediate stage is crucial for conditioning some further process. That is, the rule changing A to B feeds some other rule, which applies before B changes back into A:

(2) Feeding DY Derivation

	Underlying	/CAD/	
⇨	A→B/__D	CBD	A→B sets up environment for next rule.
	C→E/__B	EBD	Now B conditions C→E change.
⇨	B→A/__D	EAD	B→A, undoing effect of first rule.

In derivations like this, the intermediate stage is independently motivated, since it supplies the context for the change from C to E.

Vacuous DY derivations like (1) are abundantly attested; in fact, all of Pullum's examples are like this, as are many others in the literature. As the vacuity of the intermediate stage suggests, there is no need here for a serial derivation. Rather, the vacuous DY case involves blocking under constraint domination, a well-understood mode of interaction in OT (Prince and Smolensky 1993: Chapter 4). The goal of section 2 below is to demonstrate this result.

In contrast, feeding DY derivations have scarcely ever been reported in the literature, and Pullum cites no actual examples. Several possible cases are discussed and reanalyzed in section 3, with a particular concentration on the best-documented example, the interactions of syllabic and metrical processes in Bedouin Arabic. The conclusion I reach is that, in general, feeding DY derivations do not exist. This typological result demands an explanation, and in the following sections of this article I offer one.

One element of the explanation is sympathy theory (McCarthy 1999b), which is summarized in section 4. Sympathy is a general model of opaque interactions within OT. It assumes that, in addition to the actual output form, there may be a sympathetic candidate, which is the most harmonic candidate that obeys some specified faithfulness constraint. The output form is required to resemble the sympathetic candidate in some respect, and in this way the sympathetic candidate, even if not the winner itself, may exercise an indirect influence over the outcome.

The other element of the explanation is a refinement of sympathy theory, called *cumulativity*. In a DY derivation, later steps do not accumulate the results of earlier steps, since some later step literally undoes the effect of an earlier step. In non-DY derivations, later steps do reliably accumulate the mappings made earlier. A definition of cumulativity in terms of shared unfaithful mappings is proposed in section 5, and this definition is incorporated into the theory of sympathy, replacing an earlier approach based on inter-candidate faithfulness constraints. The resulting theory is one which can deal with opaque interactions generally, but which cannot accommodate the unattested feeding DY type.

The cumulativity property has implications for the theory of syllabification, and these are explored in section 6. It is not uncommon to find serial derivations in which a segment is syllabified one way, triggers some phonology, and then is resyllabified another way — a seemingly non-cumulative derivational path. The following hypothetical example is a good illustration, closely paralleling the feeding DY case in (2):

(3) Derivation with Resyllabification

	Underlying	/apia/	
⇨	Syllabification	a.pi.a	One syllable for each vocoid.
	a→i /__.	i.pi.a	Raise *a* to *i* in an open syllable.
⇨	Resyllabification	ip.ya	Resolve hiatus by devocalizing and resyllabifying.

I claim that cumulativity is defined in terms of faithfulness, and so any property that is not governed by faithfulness — arguably including syllabification — is irrelevant to determining whether a derivation is cumulative or not. Therefore, mappings like (3) are cumulative, appearances to the contrary, and so they can be simulated with the revised theory of sympathy.

2 Vacuous Duke-of-York derivations

2.1 The core cases

There is no shortage of real DY derivations of the vacuous type. Some examples, most of which were originally collected by Pullum (1976), appear in (4):

(4) Vacuous DY Cases

a. Nootka rounding/unrounding (Campbell 1973, Sapir and Swadesh 1978)

b. Vedic Sanskrit glide/vowel alternations (Kiparsky 1973a).

c. Dutch devoicing/voicing assimilation interactions (Booij 1995, Lombardi 1991).

d. English *r* deletion/intrusion (Bakovic 1998, Halle and Idsardi 1997, McCarthy 1991, 1993).

e. English trisyllabic shortening/C*i*V lengthening (Halle 1995, Kenstowicz 1994, Prince 1996).

f. Bedouin Arabic vowel raising/lowering (Al-Mozainy 1981, Irshied and Kenstowicz 1984).

g. Anglian breaking/smoothing (Dresher 1993, Hogg 1978)

These cases share certain characteristic properties: there are two (or more) rules that produce opposite mappings (A→B and B→A); these rules apply in environments that sometimes overlap; and the rules are ordered with the A→B rule applying before the B→A rule. This constellation of properties will yield a DY derivation in any word that happens to match the environment of both rules. I am using the term *vacuous* to describe these cases because the intermediate stage serves no independent function, beyond its obvious role in negotiating a path between the two contradictory rules.

Nootka nicely illustrates these observations. In Nootka, dorsal consonants (velars and uvulars) become labialized after round vowels (5a). Nootka also has underlying labiodorsal consonants, and these delabialize syllable-finally (5b). Now consider the situation where a dorsal consonant is both preceded by a round vowel and followed by a syllable boundary (indicated by “.”), so it meets the structural conditions of both rules. In fact, Delabialization takes precedence (5c):

(5) Nootka Labialization and Delabialization

a. Dorsals become labialized after round vowels

K → K^{w} / *o*___	ʔo.**k**^{w}iːɬ	‘making it’
	cf. **k**iːɬ	‘making’

b. Syllable-final labiodorsals delabialize

K^{w} → K / ___.	ɬaː**k**.šiƛ	‘to take pity on’
	cf. ɬaː.**k**^{w}iqnak	‘pitiful’

c. Interaction: Delabialization wins

	m̓oː**q**.	‘throwing off sparks’
	cf. m̓o.**q**^{w}ak	‘phosphorescent’

The problem, then, is to account for the interaction of these two processes in situations where their environments intersect.

Under the assumptions of strict serialism, the only way to ensure that Delabialization takes precedence is to order it after Labialization. The result is a DY derivation in just those cases where the ordering matters, such as input /m̓oːq/:

(6) Serial Derivation for Nootka

	Underlying	/m̓oːq/	cf.	/ʔokiːɬ/	/ɬaːk^{w}šiƛ/
⇨	Labialization	m̓oːq^{w}.		ʔo.k^{w}iːɬ	—
⇨	Delabialization	m̓oːq.		—	ɬaːk.šiƛ

Nootka, then, has exactly the characteristics of a vacuous DY derivation: two rules that produce contradictory mappings in overlapping environments are ordered so that one undoes the effect of the other.

Cases like Nootka have a straightforward non-derivational interpretation in OT, with no need for the vacuous intermediate stage. The interaction between the labialization and delabialization processes is a matter of conflicting markedness constraints, and this conflict is resolved, like all constraint conflicts, by ranking. The constraints themselves are universal; their interaction through ranking is language-particular and learned. Here I will focus on just the interaction, glossing over details of constraint formulation that are not relevant in this context.

Two markedness constraints are visibly active in Nootka. One asserts that plain dorsals cannot occur after round vowels (7a). The other prohibits rounded dorsals syllable-finally (7b):

(7) Markedness Constraints for Nootka

a. "ROUNDING"

*oK

b. "UNROUNDING"

*K^w.

These markedness constraints dominate the faithfulness constraint IDENT(round), producing the alternations in (5a, b):[2]

(8) "ROUNDING" >> IDENT(round)

/ʔokiːɬ/	"ROUNDING"	IDENT(round)
a. ☞ ʔo.k^wiːɬ		*
b. ʔo.kiːɬ	* !	

(9) "UNROUNDING" >> IDENT(round)

/ɬaːkʷši(ƛ)/	"UNROUNDING"	IDENT(round)
a. ☞ ɬaːk.ši(ƛ)		*
b. ɬaːkʷ.ši(ƛ)	* !	

This much is the basic phonology of (de-)labialization in Nootka.

Now we turn to the cases of interest, where the ranking between the two markedness constraints is decisive. If "UNROUNDING" dominates "ROUNDING", then the output will be unrounded in situations of conflict like /m̓oːq/.

(10) "UNROUNDING" >> "ROUNDING" >> IDENT(round)

/m̓oːq/	"UNROUNDING"	"ROUNDING"	IDENT(round)
a. ☞ m̓oːq.		*	
b. m̓oːqʷ.	* !		*

Obviously, there is no need for an intermediate derivational stage or kindred notion. (See Dresher (1993: 238), where a similar point is made for a mixed rule-and-constraint theory.) As usual in OT, ranking permutation predicts a range of permitted interlinguistic variation. So, if the ranking of the two markedness constraints were reversed, then *m̓oːqʷ* would be the output.

Before we continue, it is necessary to consider and dismiss two alternatives that might seem like reasonable ways to side-step the DY problem within a rule-based derivational framework. One approach, advocated by Halle and Idsardi (Halle 1995, Halle and Idsardi 1997) (cf. Bakovic 1998, Prince 1996, 1997), involves disjunctive ordering under the Elsewhere Condition (EC — Anderson 1974, Hastings 1974, Kiparsky 1973a, Koutsoudas et al. 1974, Sanders 1974). Halle and Idsardi propose to eliminate DY derivations by giving the B→A rule disjunctive precedence over the A→B rule. For example, Nootka would be analyzed by applying Delabialization before Labialization, with Labialization blocked, in EC fashion, from applying to the output of Delabialization:

(11) Nootka with Disjunctive Ordering

Underlying	/m̓oːq/[3]	cf.	/ʔokiːɬ/	/ɬaːkʷšiƛ/
Delabialization	m̓oːq.		—	ɬaːk.šiƛ
Labialization	*blocked by EC*		ʔo.kʷiːɬ	—

This proposal, if successful, would eliminate the need for the vacuous intermediate stage in DY derivations.

There is, however, a significant problem with this idea: the characteristics of DY cases are not in general the same as the characteristics of EC cases, and so the EC does not always have the desired effect. All versions of the EC require that the two rules stand in a **specific/general relation** in order for them to be disjunctively ordered. But to produce a DY derivation, the two rules only need to **overlap** in their applicability. Therefore, the conditions that trigger the EC are more stringent than the conditions that will produce a DY derivation. This means that the EC can address only a proper subset of DY derivations. Nootka illustrates this point, since Delabialization and Labialization are not in a specific/general relation. (To be at the end of a syllable is not in any way more specific than to be after a round vowel.) This observation means that the EC will not produce disjunctive application in Nootka, and so this DY case is not eliminated, nor is the more general problem solved.[4]

Another unsatisfactory approach to Nootka would involve skirting the intermediate stage of the /m̓oːq/ → *m̓oːqʷ* → *m̓oːq* derivation by enforcing the effect of Labialization in the underlying representation: /m̓oːqʷ/ → *m̓oːq*. But this means that Labialization must function as a morpheme structure constraint ruling out */m̓oːq/ and as a regular rule heteromorphemically, as in (5a). This is an instance of the Duplication Problem (Clayton 1976, Kenstowicz and Kisseberth 1977): the same rule appears twice in the grammar, in both static and dynamic roles. OT solves the Duplication Problem by denying the existence of morpheme structure constraints or other language-particular restrictions on underlying forms, deriving all linguistically significant patterns from constraints on outputs interacting with faithfulness constraints ("richness of the base" in Prince and Smolensky 1993). Because faithfulness is bottom-ranked in (10), the choice of input — /m̓oːq/, /m̓oːq^{w}/, or archisegmental /m̓oːQ/ — does not matter, since all map to surface *m̓oːq*. So there is no need to restrict the inputs and no Duplication Problem.[5]

All of the vacuous DY cases cited in (4) can be understood, like Nootka, in terms of conflict among markedness constraints resolved by ranking. The purely artifactual status of the intermediate derivational stage is revealed by this analysis. In serial theories, precedence relations among processes must be analyzed in terms of rule ordering (unless auxiliary principles like the EC intervene): the last rule to get its hands on the representation has precedence, in the sense that it reliably states a surface-true generalization. If two rules perform contradictory mappings in overlapping environments, some DY derivations are unavoidable, since there is no other way to specify the precedence relation between them. In OT, however, precedence relations among constraints are accounted for by ranking: the highest-ranking constraint has precedence, in the same sense that it reliably states a surface-true generalization.[6] By decoupling precedence from serial ordering, OT permits vacuous DY derivations to be analyzed without positing a spurious intermediate stage.

2.2 Variations

Certain other examples of DY derivations, though not strictly of the vacuous type, are also reanalyzeable in terms of conflicting constraints. Consider first the interaction of coda devoicing and voicing assimilation in Harris's (1993) analysis of Catalan:

(12) Catalan (after Harris 1993: 185f.)

From lexical stratum	sub.lu.nar	
Devoicing	sup.lu.nar	
Spirantization	*does not apply*	
Voicing Assimilation	sub.lu.nar	'sub-lunar'

This is a *bleeding* DY derivation: Devoicing bleeds Spirantization, which only affects voiced stops, but then the intermediate *p* is re-voiced by Voicing Assimilation. (Compare *su.[β]lim* 'sublime', where Spirantization applies as expected to onset *b*.)

A more direct analysis is possible, however, in terms of constraint conflict. Language typology shows that Universal Grammar (UG) contains a constraint barring continuants from codas — Korean is a well-known example where this constraint is undominated and produces alternations; similar facts in Kiowa lead Zec (1995: 111f.) to posit precisely this constraint. In Catalan, its activity is more limited: it dominates the markedness constraint responsible for the spirantization process, blocking spirantization of codas. (This is similar to Mascaró's (1984) account of Catalan.) With the markedness constraint responsible for voicing assimilation ranked above the constraint responsible for devoicing, the Catalan DY derivation reduces to vacuous status.[7]

This discussion of Catalan suggests a general approach to bleeding DY derivations, where the intermediate stage waits out a third rule. The general form of such derivations is this:

(13) Bleeding DY Derivation

	Underlying	/CAD/	
⇨	A→B/__D	CBD	A→B to escape next rule
	C→E/__A	*does not apply*	No A there to condition C→E change.
⇨	B→A/__D	CAD	B→A, undoing effect of first rule.

Descriptively, the effect of this derivation is to change /C/ into *E* before *A*, except when *A* is followed by *D*. This is just the familiar blocking pattern obtained by ranking markedness constraints, as in Nootka or Catalan. The constraint *CA can compel the unfaithful mapping of /C/ to *E*, but *CA is crucially dominated by another markedness constraint, *EAD, which effectively blocks that mapping. So, in the general case, bleeding DY derivations can be reduced to vacuous status.[8]

Another variation on the DY theme can be found in Rubach's (1993: 266ff.) analysis of depalatalization in Slovak (or Polish (Rubach 1984: 101ff., 199f.)). The rule of Anterior Depalatalization affects palatalized *t′, d′,* and *n′* when they precede coronals: *kost′* 'bone', *kostný* 'bony'. In addition, Anterior Depalatalization "undoes the effect of Coronal Palatalization whenever Coronal Palatalization has applied before a yer-initial suffix containing a coronal consonant and the yer has not been vocalized" (Rubach 1993: 267):

(14) Palatalization and Depalatalization in Slovak

Underlying	/let+En+ý/	
Palatalization	let′Ený	
Yer Vocalization	*does not apply*	
Yer Deletion	let′ný	
Depalatalization	letný	'summery'

The yers *E* and *O* are abstract vowels posited in most analyses of Slavic languages. When followed by another yer in the next syllable, a yer "vocalizes" to *e* or *o*; otherwise, as in (14), it deletes. Before it deletes, the front yer *E* causes palatalization of a preceding consonant (cf. *mliek*/*mliečný* 'milk'/'milky'). Once the yer has deleted, though, the *t′* in (14) is followed by a coronal, and so it must depalatalize.

This too is a vacuous DY derivation, though with the added complication of an opaque interaction between Palatalization and Yer Deletion. In OT, the opaque interaction can be accounted for using sympathy theory (as in Łubowicz's (1999) analysis of similar facts in Polish). The sympathy constraint is, however, crucially dominated by a constraint against clusters like *t'n*. Schematically, except for the opacity, the interaction here is no different than Nootka or Catalan.

2.3 Summary

I have now reviewed the case of vacuous DY derivation and variations on it. I have shown that the vacuous DY pattern is an expected consequence of the core premises of OT, constraint ranking and constraint violation under domination. Significantly, cases of this type are well-attested and uncontroversial, indicating that the typological claim implicit in OT fits the facts. But when we turn to the feeding DY interaction in the next section, the situation is quite different.

3 Feeding Duke-of-York derivations

3.1 Introduction

In feeding DY derivations like (2), repeated immediately below, the intermediate stage is crucial:

(2) Feeding DY Derivation

	Underlying	/CAD/	
⇨	A→B/__D	CBD	A→B sets up environment for next rule.
	C→E/__B	EBD	Now B conditions C→E change.
⇨	B→A/__D	EAD	B→A, undoing effect of first rule.

The A→B rule feeds some process which applies at the intermediate stage, before the B→A rule wipes out its environment.

Plausible-looking examples are not difficult to concoct. The first, given in (15), is modeled after a post-vocalic spirantization process in Tiberian Hebrew, but with a twist. In this hypothetical case, rules epenthesizing and later deleting ə are wrapped around a process of post-vocalic spirantization:

(15) Quasi-Hebrew (Hypothetical Feeding DY Derivation)

	Underlying	/qarbi/	
⇨	Epenthesis	qarəbi	Insert ə after any syllable coda.
	Spirantization	qarəvi	Stops become fricatives post-vocalically.
⇨	Syncope	qarvi	Delete ə in two-sided open syllable (VC__CV)

In feeding DY fashion, the ə is inserted, hangs around long enough to cause spirantization, and then deletes, leaving a fricative behind as evidence of its passage.

A more complex hypothetical example was brought to my attention by Paul Kiparsky (e-mail communication, 7 July 1998). At the first step, trimoraic CVːC syllables are repaired by *i* epenthesis (cf. Mekkan Arabic in Abu-Mansour 1987). The vowel *i*, whether underlying or epenthetic, then triggers palatalization of a preceding coronal. A process of apocope deletes final vowels, including epenthetic *i*, and finally the CVːC syllable is re-repaired by shortening. Because it shares some rules with the real Yokuts language, I will call this hypothetical system quasi-Yokuts:

(16) Quasi-Yokuts (Hypothetical Feeding DY Derivation)

	Underlying	/maːt/	
⇨	Epenthesis	maːti	To repair trimoraic syllable.
	Palatalization	maːči	*ti* → *či* generally.
⇨	Apocope	maːč	Final vowels delete
	Shortening	mač	To repair trimoraic syllable.

The vowel *i* is epenthesized, triggers palatalization, and later deletes. The conditions that originally produced *i* epenthesis, a trimoraic syllable, is subsequently repaired by other means.

In quasi-Hebrew and quasi-Yokuts, some crucial phonological business occurs at the intermediate stage of the derivation — unlike the vacuous DY cases of section 2. In quasi-Hebrew, the intermediate stage allows the temporary ə to condition post-vocalic spirantization, and in quasi-Yokuts, the intermediate stage is the point at which temporary final *i* triggers palatalization. What makes these cases particularly interesting is that each process individually is quite plausible and natural. The peculiar thing is not the rules themselves, but their co-existence and interaction in a single system.

3.2 Review of putative examples

Examples of feeding DY derivations are not exactly thick on the ground, or even thin. Pullum's (1976) survey contains none, and I am aware of just four putative cases:

(i) Insertion and removal of coda moras in Tübatulabal (Crowhurst 1991).

(ii) Harmony and disharmony of neutral vowels in, e.g., Finnish (Bach 1968 et seq.).

(iii) Epenthesis and syncope of ə in (real) Tiberian Hebrew (Prince 1975).

(iv) Syncope and epenthesis in Bedouin Arabic (Al-Mozainy 1981).

I will pass over (i) and (ii) fairly quickly, since there are equally good and possibly superior alternatives to the DY derivations. I will then show that Hebrew involves a cyclic or "output-output" faithfulness effect. Finally, I will turn to a close examination of the Bedouin Arabic case, asking whether it is an authentic instance of the feeding DY type or not. I will argue that it is not, concluding that feeding DY interactions do not in general occur — an observation for which linguistic theory needs to supply an explanation.

(i) Coda moras in Tübatulabal. Crowhurst (1991) argues that an early rule, Reduplication, treats CVC syllables as bimoraic, while a later rule, Stress Assignment, treats CVC syllables as monomoraic. This requires a feeding DY derivation: codas are assigned a

mora by Weight-by-Position (Hayes 1989), Reduplication applies, coda moras are deleted, and then Stress Assignment applies. Reduplication crucially relies on the intermediate stage where coda moras are temporarily present.

This chain of reasoning relies on the assumption (which Crowhurst shares with several contributors to this volume) that stress assignment is reliably diagnostic of the mora count of CVC syllables. That assumption has been called into question in recent years. Research on syllable weight (e.g., Gordon 1999, Hayes 1995, de Lacy 1997) has shown that stress is also conditioned by factors like sonority that are not reified in the moraic representation. Arguably, Tübatulabal is just such a case: stress is attracted to certain syllables because they contain prominent long vowels, not because they are the only bimoraic syllables at the derivational instant when stress is assigned. On the strength of the reduplicative evidence, then, and with no remaining impediments from the stress evidence, it is reasonable to suppose that CVC syllables in Tübatulabal are bimoraic *tout court*.

(ii) Neutral vowels in harmony systems. Ever since Bach (1968), a common analytic strategy for dealing with neutral vowels has been to assume that they temporarily undergo the harmony process: e.g., Finnish /tuoli-llA/ → *tuol**ɨ**lla* → *tuol**i**lla* ‘on the chair’. (See Ní Chiosáin and Padgett (1997) and Walker (1998) for a similar approach within OT.) This too is a feeding DY interaction, since the intermediate stage is required to support a strictly local, iterative harmony process.

This idea is not lightly dismissed, though it is worth noting that the principal motivation for strict locality in current thinking (see also Archangeli and Pulleyblank 1994, Gafos 1996, 1999, Gafos 1998, Pulleyblank 1996), a kind of phonetic realism, is difficult to reconcile with the fact that the phonetically real representations do not actually respect strict locality. There are alternatives to strict locality, summarized with references in Bakovic (2000: 266ff.), including an approach, proposed by Bakovic, that is strictly local but without

DY derivations, based on an extension of Wilson's (1999) "targeted constraints".

(iii) Spirantization in Tiberian Hebrew. A process of post-vocalic spirantization in Hebrew is rendered opaque by syncope: /katab(+uː)/ → *kaːθaβ*/*kaːθβuː* 'he/they wrote'. According to Prince (1975), there is one particular morphological situation where a vowel is inserted, remains around long enough to trigger spirantization, and then is syncopated. This happens when the infinitival stem /ktob/ bears a prepositional prefix like /bi#/ 'in':[9]

(17) Tiberian Hebrew Feeding DY Derivation (after Prince 1975)

	Underlying	/bi#ktob/
⇨	Epenthesis	bi#kətob
	Spirantization	bi#xəθoβ
⇨	Syncope	bi#xθoβ
	Other rules	bixθoːβ

Unlike the invented example in (15), though, real Hebrew has no general process epenthesizing schwa in a context that will later trigger syncope. In fact, the infinitive *lix***t***oːβ* 'to write' supplies a near-minimal pair, with the expected stop *t* and no ghost of a prior epenthesis process. Derivations like (17) are limited to words that bear the syntactically independent prefixes /bi#/ and /ki#/ 'like', as Prince indicates with the # boundary. With ordinary inflectional prefixes (*lix***t***oːβ*) or morpheme-internally (*mal***k***iː* 'my king'), there is no Ø → ə → Ø DY derivation.

The morphology is obviously the key to understanding this restricted DY effect. In the theory of Lexical Phonology, one would say that Epenthesis and Spirantization apply to /ktob/ at a stratum when /bi/ is not yet present, producing the free-standing word *kə.θoːβ* 'writing'. Only later is /bi/ added, triggering syncope of ə. Words with inflectional prefixes like *lixtoːβ* and tautomorphemic cases like *malkiː* are derived in the earlier stratum, so Epenthesis never applies.

In OT, an approximation to the Lexical Phonology analysis is possible using output-output correspondence. The spirantized *θ* in *bixθoːβ* is faithful to its correspondent in the free-standing word *kə.θoːβ* in obedience to OO-IDENT(cont). The difference in strata is modeled by allowing different affixes to assign different output-output correspondence relations (Benua 1997a, 1997b). In this way, a restricted feeding DY effect — limited to circumstances where the intermediate stage is another independent word — can and should be reconstructed in OT. Of course, standard serial phonology is subject to no such restriction; it allows feeding DY derivations even morpheme-internally, where there is no evidence for cyclic or stratal organization.

3.3 A feeding Duke-of-York interaction in Bedouin Arabic

The standard analysis of Bedouin Arabic incorporates a feeding DY interaction between stress and syllabically-conditioned rules of vowel deletion and epenthesis. Words like /ʔakal-at/ 'she ate' are said to get initial stress *ʔákalat*, followed by deletion of the stressed vowel with concomitant shift of the stress to the following syllable *ʔkálat*, and later epenthesis to restore the deleted vowel *ʔakálat*. As we will see below (3.3.1), the evidence in support of this DY derivation is quite compelling, and it is not lightly dismissed.

Nonetheless, there is a better analysis that avoids the need for the stress→deletion→stress-shift→epenthesis DY derivation. The problem in traditional accounts lies with the deletion rule, which purportedly deletes a vowel in a light syllable that is itself followed by a non-final light syllable (3.3.2). This rule's complex, non-local environment amounts to nothing more than a re-description of the facts. But an explanation is possible in terms of prosodic theory. This deletion process, I will argue, is an instance of the well-documented tendency of iambic feet to maximize quantitative differences between their head and dependent syllables. This analysis, in common with several other contributions to this volume, relies on positing mora-less semi-syllables in the output of vowel deletion. The analysis explains all properties of the deletion process, and it eliminates the need for the DY derivation.

In section 3.3.3, this analysis is formalized within OT. One aspect of stress/deletion interaction turns out to involve phonological opacity. Opacity is intimately connected with the DY problem, because DY derivations are by their very nature opaque (pace Pullum 1976: 89-90). The following sections (4–6) then go on to address the general problem of opacity and DY in OT.

3.3.1 Overview of traditional analysis. To understand the DY derivation, it is first necessary to have a good deal of background in Bedouin Arabic phonology. The plan is first to present the core processes of vowel raising and deletion, and then turn to their interaction with stress, which has DY character.

The following partial paradigms illustrate the main points:[10]

(18) Bedouin Arabic Core Data

	/katab/ 'wrote'	/samiʕ/ 'heard'	/kitib/ 'was written'
'he___' +Ø	kítab	símiʕ	ktíb
'we___' +*na*	kitábna	simíʕna	ktíbna
'she___' +*at*	ktíbat	sámʕat	kítbat

Descriptively, underlying /a/ raises to *i* in an open syllable, while underlying /i/ deletes in the same environment — a typical chain-shift. But observe that even underlying /a/ has deleted in the form *k_tíbat* (from /katab–at/).[11]

Starting with Al-Mozainy (1976, 1981) and continuing through Al-Mozainy et al. (1985), Hayes (1995), and Irshied and Kenstowicz (1984), most analysts have agreed on approximately the following rule system to deal with the data in (18):

(19) Core Rules for Bedouin Arabic

a. Syncope

i → Ø / __.σ   Delete short *i* in a non-final light syllable.

b. Trisyllabic Deletion

V → Ø /___. L σ   Delete a short vowel from an open syllable that is followed by a non-final light syllable.

c. Raising

a → i / ___. σ   Raise *a* to *i* in a non-final open syllable.

The rule of Syncope is necessary to account for alternations like *samiʕ/sam_ʕat*. Raising is exemplified by forms like *k**i**tab*. Trisyllabic Deletion will be discussed in detail below (3.3.2).

These rules have several crucial ordering relations which are illustrated by the following derivations:

(20) Rule Interaction

Underlying	/katab/	/katab-at/	/samiʕ/	/samiʕ-at/	/kitib/	/kitib-at/
Initial Syllabification	ka.tab	ka.ta.bat	sa.miʕ	sa.mi.ʕat	ki.tib	ki.ti.bat
Syncope i → Ø / C__ σ				sam.ʕat	ktib	kit.bat
Trisyllabic Deletion V → Ø / C__ L σ		kta.bat				
Raising a → i / C__ σ	ki.tab	kti.bat	si.miʕ			

After an initial round of syllabification, Syncope first applies, deleting all *i*'s that occur in open syllables. (To handle /kitib-at/, right-to-left iteration of Syncope has to be assumed.) Syncope crucially precedes Trisyllabic Deletion, since otherwise the first vowel of /samiʕ-at/ would be deleted. Syncope must also precede Raising, since otherwise the first vowel of /samiʕ-at/ would raise to *i*. This ordering — Syncope, then Raising — is responsible for the

/a/ → *i*, /i/ → Ø chain shift which can be observed in these examples.

The interaction of Trisyllabic Deletion with Stress is the source of the DY derivation. Standard accounts posit a Latin-type stress rule, as in the sedentary Arabic dialects discussed by Kiparsky (this volume) and Wiltshire (this volume): stress the penult if heavy, otherwise the antepenult.[12] Formally, a moraic trochee is assigned right-to-left, subject to extrametricality of the final syllable: *kítab*, *sámʕat*, *kitábna*, *maktúːʃah* 'tied (f. sg.)', *máːlana* 'our property', *ð̣arábtukum* 'I hit you (m. pl.)', *yšuːfúːnukum* 'they (m.) see you (f. pl.)'. But in words that are subject to Trisyllabic Deletion, the traditional analysis posits an early stress rule followed by stress shift when the stressed vowel deletes (Al-Mozainy et al. 1985, Al-Mozainy 1981, Hayes 1995: 228-38):

(21) Interaction of Stress and Trisyllabic Deletion in Standard Analysis

		a. /katab-at/	b. /ʔinkasar-at/	c. /ʔakal-at/
	Stress	(káta)bat	ʔin(kása)rat	(ʔáka)lat
⇨	Trisyllabic Deletion	(kta)bat	ʔin(ksa)rat	(ʔka)lat
	Stress Shift	(ktá)bat	ʔin(ksá)rat	(ʔká)lat
⇨	Epenthesis (#ʔ_C)			(ʔaká)lat
	Other rules	ktíbat	ʔinksárat	ʔakálat

Deletion of a vowel out from under the stress forces stress to shift to the other syllable of the foot. The derivation in (21a) is provided for comparison purposes; the interesting cases are (21b) and (21c). In (21b), Latin-type trochaic stress is applied at an early stage of the derivation, but then it is obscured by the subsequent effects of Trisyllabic Deletion and concomitant stress shift. In (21c), this much also happens, plus the effect of Trisyllabic Deletion is undone, in classic DY fashion, by an epenthesis rule which repairs the initial #ʔC cluster. This is a perfect exemplar of a feeding DY derivation, because the intermediate stage is crucial to obtaining the stress-shift effect.

3.3.2 The prosodic basis of Trisyllabic Deletion. There is good reason to be skeptical of Trisyllabic Deletion and the DY derivation based on it. Trisyllabic Deletion has a complex, non-local, and highly arbitrary environment — why should deletion be limited to a light syllable that is followed by a light syllable that is itself non-final? The conditioning factors don't seem to make sense.

This situation is strongly reminiscent of trisyllabic shortening in English. Pairs like *ser**e**ne/ser**e**nity*, *gr**a**teful/gr**a**titude*, and *der**i**ve/der**i**vative* show that, descriptively, a long vowel is shortened when followed by an unstressed syllable that is itself non-final. The standard analysis (Chomsky and Halle 1968) uses a rule with a complex, non-local environment, much like Trisyllabic Deletion. Again, the conditiong factors don't seem to make sense.

The explanation for the English case (and its Arabic counterpart) comes from higher-level prosodic structure. According to Myers (1987) and Prince (1990), trisyllabic shortening is conditioned by foot structure, in top-down fashion. The typical English pattern is a trochaic foot over penult and antepenult, with final syllable extrametricality: *se (réni)*$_{Ft}$ *⟨ty⟩*. Shortening improves the well-formedness of the trochaic foot, replacing a HL (heavy-light) trochee with a more harmonic LL trochee (Prince 1990). This approach has answers to the *why*'s of trisyllabic shortening, as Prince (1996) emphasizes. Why shortening and not, say, lengthening? Because shortening improves the match with the preferred bimoraic foot. Why a following unstressed syllable? Because a following unstressed syllable is a descriptive artifact of the real foot-based condition. And why, in non-local fashion, must there be another syllable after that? Because of the regular extrametricality rule. The answers to these questions emerge once the role of trochaic foot structure in English is properly understood, while they remain mysteries under the standard analysis.

Similarly, I propose that the key to understanding trisyllabic deletion in Bedouin

Arabic is placing it in the context of an *iambic* stress system, contrary to the standard trochaic analysis. Iambic feet are subject to strong quantitative requirements. According to the Iambic/Trochaic Law (Hayes 1987, 1995, 80) or Grouping Harmony (Prince 1990) (cf. also McCarthy and Prince 1986), iambic feet tend to favor quantitative reinforcement of the prominential contrast, so a LH iamb is better than a LL one. For concreteness, I will assume the following formulation of Grouping Harmony:

(22) GRPHARM

In an iambic foot (x 'y), |y| > |x|. (|α| ≡ weight of α in moras)

Because of GRPHARM, many languages have iambic lengthening processes, which improve LL iambs by lengthening the second syllable. Another logically possible consequence of GRPHARM is reduction of unstressed syllables in iambic feet, enhancing the quantitative contrast by weakening the weak rather than strengthening the strong. Hayes (1995, 213) reports that this occurs in Delaware, and it is an element of Kager's (1997) analysis of Macushi Carib. Trisyllabic deletion, I will show, is exactly this: reduction of the unstressed syllable in a LL iambic foot to enhance the quantitative contrast.

First, though, we must establish that the stress system of Bedouin Arabic is indeed iambic. Traditionally, the Arabian Bedouin Arabic dialects have been assumed to have trochaic stress, like all the sedentary dialects, and Al-Mozainy, among others, adopts that assumption. But Hayes (1995) shows that two non-Arabian Bedouin dialects, one spoken in eastern Libya and the other in the Negev, are actually iambic. I will now show that Al-Mozainy's Arabian dialect is also iambic.

The analysis of words like *ʔakálat* is tortuous under trochaic assumptions, but if stress is left-to-right iambic, then the analysis is straightforward: *(ʔaká)lat*. Moreover, the examples cited above in support of the trochaic analysis — *kítab*, *sámʕat*, *kitábna, maktúːfah*, *máːlana*, *ḍarábtukum*, *yšuːfúːnukum* — are also compatible with left-to-right iambic feet. In words with heavy penults like *kitábna*, trochaic footing (*ki(táb)na*) and iambic footing (*(kitáb)na*) produce descriptively equivalent results. In disyllables like *kítab*, iambic FT-FORM yields to NON-FINALITY, which is a nea r-universal accompaniment to iambic stress (Hung 1994, Prince and Smolensky 1993). Thus, disyllables do indeed have trochaic stress — as in *(kítab)* — but only when higher-ranking NON-FINALITY compels violation of FT-FORM(IAMBIC). Words with heavy antepenults and light penults, such as *ð̣arábtukum*, follow the "foot extrametricality" pattern identified by Hayes (1995: 232). The actual output form is *(ð̣aráb)tukum*, and its most important competing candidate is **ð̣arab(túkum),* which also satisfies NON-FINALITY but violates FT-FORM. In contrast, the actual output form satisfies NON-FINALITY and FT-FORM at the price of inferior rightward alignment of its main stress (ALIGN-HEAD-R (McCarthy and Prince 1993a)). We therefore have the ranking NON-FINALITY >> FT-FORM >> ALIGN-HEAD-R, as the following tableau certifies:

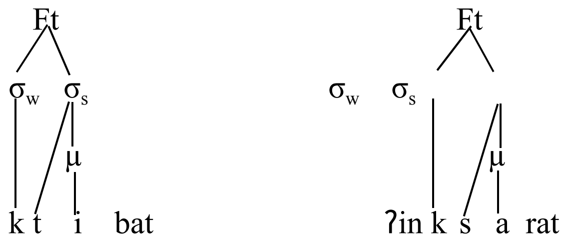
(23) Iambic Stress in *(ð̣aráb)tukum*

	NON-FINALITY	FT-FORM	ALIGN-HEAD-R
a. ☞ (ð̣aráb)tukum			**
b. ð̣arab(túkum)		*!	*
c. ð̣a(rábtu)kum		*!	**
d. ð̣arab(tukúm)	*!		

The two threads of analysis, GRPHARM and iambic stress, can now be combined to supply an explanation for the trisyllabic deletion process. Without trisyllabic deletion, a word like /ʔinkasar-at/ would be parsed with a LL iambic foot: **ʔin(kisá)rat*. Trisyllabic deletion improves the quantitative structure of this iamb. According to GRPHARM, iambic feet optimally match their weak-strong prominence with short-long quantity. Many languages have iambic lengthening, where a LL iambic foot becomes LH by lengthening the vowel of the second syllable. In Bedouin Arabic, I claim, a LL iamb becomes ΔL, where Δ denotes a moraless syllable, called a *semi-syllable* (see below):

(24) The ΔL Iamb in Trisyllabic Deletion Cases

a. ktíbat b. ʔinksárat



The idea, then, is that loss of the pre-stress vowel in *ʔin(ksá)rat* brings this word into conformity with GrpHarm, in a way that closely parallels iambic lengthening effects in other languages.

There are several reasons to think that this account of trisyllabic deletion is essentially correct.

First, it offers a complete, strictly local explanation for the peculiar contextual conditions on trisyllabic deletion:

(25)	Observation	Explanation
a.	Trisyllabic deletion only affects a light syllable.	Only a light syllable can be the weak branch of an iambic foot.
b.	The affected syllable must be followed by another light syllable.	If the following syllable is heavy, then the iamb is already LH, satisfying GrpHarm without further ado: *(ki.táb)na*, *(ḍaráb)tukum.*
c.	The syllable following the affected syllable must itself be non-final.	If the following syllable is final, then the foot is trochaic, not iambic, because Non-Finality dominates Ft-Form (23): /rama/ → *(ríma)* ‘he threw’.

Second, this analysis explains a significant correlation in the history of Arabic dialects. The sedentary dialects have trochaic stress, and they never have trisyllabic deletion. The Bedouin dialects have iambic stress, and many (though not all) have trisyllabic deletion. Historically, then, trisyllabic deletion appears to be a secondary development in those dialects that first changed to iambic stress — exactly as the synchronic analysis predicts.

Third, this analysis also accords well with processes affecting iambic feet in other languages, as documented by Hayes (1995) and Kager (1997). Kager's analysis of Macushi Carib is a close parallel in many respects.

Fourth, this analysis makes sense syllabically. What appear to be tautosyllabic clusters arise only as a result of vowel deletion, supporting the claim that they actually involve semi-syllables : *(k.tí).bat*, *ʔin.(k.sá).rat*.[13] This too is closely paralleled in Kager's analysis of Macushi Carib. And overall, there is ample precedent for semisyllables or similar notions in Arabic (Kiparsky, this volume, and Aoun 1979, Broselow 1992, Farwaneh 1995, McCarthy and Prince 1990a, 1990b, Selkirk 1981), in other languages (e.g., Cho and King, this volume, Féry, this volume), and in analyses of epenthesis (Hyman 1985, Piggott 1995).

Finally, this analysis accounts for words like *ʔinksárat* without the problematic stress-shift process. There is instead iambic stress, with optimization of the quantitative relations in the iambic foot.

In short, trisyllabic deletion is actually iambic deletion — a local process, motivated by foot well-formedness, much like the Myers-Prince approach to trisyllabic shortening in English.

3.3.3 OT Analysis of Bedouin Arabic. These ideas can be incorporated into a fuller OT analysis of Bedouin Arabic, which also deals with the reduction and syncope processes. The first order of analytic business is to dispose of the /a/ → *i*, /i/ → Ø chain shift. The insight behind the analysis of chain shifts in OT is relative faithfulness (Gnanadesikan 1997, Kirchner 1996): if /A/ → B and /B/ → C in the same environment, but /A/ ✗ C, then the prohibited /A/ → C mapping must be categorically less faithful than the permitted /A/ → B and /B/ → C mappings.[14] Then the markedness constraint that drives these alternations can be ranked so that it can compel the "shorter" mappings but not the "longer" one.

Gnanadesikan (1997) proposes that this distinction in relative faithfulness is defined on universal phonological scales, such as consonantal stricture, voicing and sonorancy, or vowel height.[15] All scales are ternary, by her hypothesis, and positions on a scale can be referenced by markedness and faithfulness constraints. The key is to recognize two distinct faithfulness constraints, one that precludes any movement on a scale and another that prohibits only longer movements on the scale:

(26) STAY(S) and STAY-ADJ(S)

a. STAY(S)

Input and output have the same position on the scale S.

b. STAY-ADJ(S)

Input and output have the same or adjacent positions on the scale S.

If there is a universal phonological scale S = A > B > C, then the mappings /A/ → B, /B/ → C, and /A/ → C all incur violations of STAY(S). But the mapping /A/ → C also incurs a violation of STAY-ADJ(S), and so it is categorically less faithful than the other two mappings.

Following Kirchner (1996), I assume that the scalar dimension relevant to the Bedouin Arabic *a*/*i*/Ø alternations is intrinsic duration (Lehiste 1970): the low vowel is longest, the high vowel is intermediate, and of course Ø is shortest:

(27) The Duration Scale *Dur*

a > *i* > Ø

Kirchner proposes that the markedness constraint driving the chain shift is REDUCE:

(28) REDUCE (after Kirchner 1996, 347)

A short vowel in an open syllable has zero duration. Assign one violation mark for each increment of duration above Ø on the scale *Dur*.

So the vowel *i* receives one mark from REDUCE, while *a* gets two. Ranked between STAY-ADJ(Dur) and STAY(Dur), REDUCE is responsible for the /a/ → *i* and /i/ → Ø chain shift, and it is correctly unable to compel the /a/ → Ø mapping:

(29) Raising: /a/ → *i* in open syllable

/katab/	STAY-ADJ(Dur)	REDUCE	STAY(Dur)
a. ☞ ki.tab		*	*
b. ka.tab		** !	
c. k.tab	* !		*

(30) Syncope: /i/ → Ø in open syllable

/kitib/	STAY-ADJ(Dur)	REDUCE	STAY(Dur)
a. ☞ k.tib			*
b. ki.tib		* !	
c. ka.tib		** !	*

In (29), the candidate with raising of /a/ to *i* triumphs over the faithful candidate by virtue of its better performance on the markedness constraint REDUCE. Perfect performance on REDUCE is available from the remaining candidate, (29c), but the cost is too high: fatal violation of top-ranked STAY-ADJ(Dur), which bars the /a/ → Ø mapping. In (30), on the other hand, perfect performance on REDUCE is possible: the /i/ → Ø mapping only violates the low-ranking faithfulness constraint STAY(Dur), since *i* and Ø are adjacent on the Dur scale. In this way, the shorter mappings (/a/ → *i*, /i/ → Ø) are permitted, but the longer mapping (/a/ → Ø) is not.

Of course, /a/ does delete when GRPHARM is at stake. To compel deletion of underlying /a/, GRPHARM must therefore be ranked above STAY-ADJ(Dur):

(31) Application to Trisyllabic Deletion Case

/katab-at/	GRPHARM	STAY-ADJ(Dur)	REDUCE	STAY(Dur)
a. ☞ (k.tí).bat		*	*	**
b. (ki.tí).bat	*!		**	**

The failed candidate (31b) contains a LL iambic foot, violating GRPHARM. The alternative in (31a) contains an iamb of properly unequal weight, obtained by deleting the first vowel, leaving only a weightless semi-syllable behind. Alternative candidates like **(ktí).bat* (with a complex onset), **(tí).bat* (with consonant deletion), or **(kitáa).bat* (with lengthening instead of shortening) violate undominated constraints, so they need not distract us further.

The DY case *(ʔaká)lat* is analyzed in much the same way, except that it shows the effect of an undominated constraint against a semi-syllable with onset *ʔ*. (This replaces the special post-*ʔ* epenthesis rule of the traditional analysis.) Ranked above GRPHARM, that constraint effectively blocks trisyllabic deletion in words with initial *ʔ*.

(32) Application to /ʔakal–at/

/ʔakal–at/	*ʔΔ	GRPHARM	STAY-ADJ(Dur)
a. ☞ (ʔa.ká).lat		*	
b. (ʔ.ká).lat	*!		*

In short, this DY case is analyzed in terms of conflicting markedness constraints, just like the vacuous DY examples discussed in section 2. There is no need for stress shift under deletion; stress is iambic in conformity with the general pattern of the language.

We now have a reasonably complete picture of the trisyllabic deletion phenomenon. Trisyllabic deletion can be explained in terms of known quantitative properties of iambic stress systems. There is no evidence for a DY derivation; instead, there is a blocking effect by virtue of one markedness constraint dominating another, as in Nootka (10).

One detail remains, and it introduces the issue of opacity, which intersects in important ways with the analysis of DY derivations. Consider the effect of adding the candidate **(kát).bat* to the tableau (31):

(33) Tableau (31) with **(kát).bat* Added

	/katab-at/	GrpHarm	Stay-Adj(Dur)	Reduce	Stay(Dur)
Opaque	a. ☞ (k.tí).bat		*	*!	**
	b. (ki.tí).bat	*!		**	**
Transparent	c. ☜ (kát).bat		*		*

This additional candidate harmonically bounds the intended output *(k.tí).bat*, a problematic condition I have indicated with the reversed pointing hand.[16] To get the right result here, there must be some further constraint, ranked above Reduce, that *(k.tí).bat* satisfies better than **(kát).bat* does.[17]

This is a case of *opacity*. Two phonological processes interact opaquely if one hides the results or environment of the other:

(34) Opacity (after Kiparsky 1973b)

A phonological rule ℘ of the form $A \rightarrow B / C\_\_\_D$ is *opaque* if there are surface structures with any of the following characteristics:

a. instances of A in the environment $C\_\_\_D$.

b. instances of B derived by ℘ that occur in environments other than $C\_\_\_D$.

Intuitively, the idea is that a rule is opaque if there are surface forms that look like they should have undergone it but did not (34a) or surface forms that underwent the rule but look like they should not have (34b). In rule-based phonology, the output of Trisyllabic Deletion, *(k.tí).bat*, is opaque with respect to Syncope, because it contains surface *i* in an open syllable

((34a)-type opacity). In Optimality Theory, the hallmark of opacity is unexplained markedness or faithfulness violation by the actual output form (McCarthy 1999b). In (33), as was already noted, the intended output *(k.tí).bat* has unexplained violations of both the markedness constraint REDUCE and the faithfulness constraint STAY(Dur). These violations are unexplained because there is another candidate, **(kát).bat*, that fares better on both of these constraints and equally well on all higher-ranking constraints. Opaque interactions demand some revision of the basic theory, and that is the subject of the next section.

4 Sympathy and opacity

The problem identified in (33) is that the actual output form *(k.tí).bat* has all of the violation marks of the failed candidate **(kát).bat*, and more. Some higher-ranking constraint must compel these violations. According to sympathy theory (McCarthy 1999b), the responsible constraint is one that is sensitive to relations between candidates — specifically, the relation between all other candidates and one particular candidate, called the sympathetic candidate (which is notated with the symbol ❀). The sympathetic candidate is chosen by faithfulness to the input: it is the most harmonic candidate that obeys some designated faithfulness constraint, called the selector (which is notated by the symbol ✯). A ranked, violable sympathy constraint (also notated by ❀) assesses candidates for their similarity, in a sense to be made precise below, to the sympathetic candidate. A sympathy constraint is responsible for compelling *(k.tí).bat*'s seemingly supererogatory constraint violations.

Even without the details of how the sympathy constraint works, we can still get a reasonably good picture of sympathy theory in action. A little bit of the logic of sympathy starts the ball rolling. If the effects of sympathy are to be non-vacuous, the sympathetic candidate must be distinct from both the actual output *(k.tí).bat* and its transparent competitor *(kát).bat*. And since the sympathetic candidate is chosen for obedience to a faithfulness constraint (the selector), it follows that it must be more faithful, on some dimension, than either *(k.tí).bat* or **(kát).bat*. This reasoning leads to STAY-ADJ(Dur) as the selector, since it is the only faithfulness constraint violated by both *(k.tí).bat* and **(kát).bat*. The most harmonic candidate that obeys ✯STAY-ADJ(Dur) is ❀*(ki.tí).bat*. It obeys the selector because no /a/'s have been deleted. It is the most harmonic candidate, given this restriction, because the /a/'s in open syllables have reduced to *i* (maximally satisfying REDUCE — compare faithful **(ka.tá).bat*) and stress is iambic.

The sympathy constraint, here temporarily designated by ❀SYM, evaluates candidates for similarity to the sympathetic candidate (in a way that will be made precise below). The actual output form *(k.tí).bat* (35a) is more similar to ❀*(ki.tí).bat* than **(kát).bat* (35b) is, and so *(k.tí).bat* performs better on ❀SYM. Obviously, ❀*(ki.tí).bat* is maximally similar to itself, and so it performs perfectly on ❀SYM, but it is not optimal because of its fatal GRPHARM violation. The following tableau adds the sympathy constraint to (33):

(35) Sympathy Applied to /katab-at/ → *(k.tí).bat*, **(kát).bat*

	/katab-at/	GRPHARM	❀SYM	✯STAY-ADJ(Dur)	REDUCE	STAY(Dur)
Opaque	a. ☞ (k.tí).bat		*	*	*	**
Sympathetic	b. ❀ (ki.tí).bat	*!		✓	**	**
Transparent	c. ☜ (kát).bat		**!	*		*
Faithful	d. (ka.tá).bat	*!		✓	****	

The numbers of violation-marks in the ❀SYM column should not be taken literally, but the relative harmony of candidates with respect to this constraint should be. It is **(kát).bat*'s inferior resemblance to the sympathetic candidate that explains why it is not optimal, thereby accounting for *(k.tí).bat*'s otherwise unexplained violations of REDUCE and STAY(Dur).[18]

To complete this sketch, we need to check that sympathy has no untoward consequences for the rest of the language. No effects of sympathy are expected if the actual output form obeys the selector constraint, because in that case the selector and normal harmonic evaluation will converge on the same candidate, and so ❀SYM will be vacuously satisfied by a candidate that would have been optimal in any case. Some perusal of the core data in (18) shows that deletion of /a/, which translates into violation of ✯STAY-ADJ(Dur), only occurs in the /katab-at/ → *(k.tí).bat* derivation, and so that is the only circumstance where sympathy is relevant.[19]

5 Sympathy and cumulativity

5.1 The problem

The issue to be addressed now is the nature of the sympathy constraint ❀SYM. In the earliest, unpublished work on sympathy theory (McCarthy 1998), the role of ❀SYM is fulfilled by a family of inter-candidate faithfulness constraints, specifying the exact way in which the candidate under evaluation must match the sympathetic candidate. For instance, inter-candidate faithfulness constraints requiring corresponding vowels to match in height or stress would correctly favor *(k.tí).bat* over **(kát).bat* in (35), since the former better matches ❀*(ki.tí).bat*'s vowel height and stress than the latter does.

This framework of sympathetic inter-candidate faithfulness constraints is very rich, because it brings with it the full expressive power of correspondence theory (McCarthy and Prince 1995). In fact, it is too rich, because it permits unattested patterns of opacity to be described, such as the feeding DY type. In this section, I will argue against inter-candidate faithfulness constraints and in favor of an alternative that is based on comparing the unfaithful mappings that produce candidates.

Perhaps the clearest example of the excessive descriptive power of inter-candidate faithfulness is the quasi-Yokuts example (16), repeated here for convenience:

(16) Quasi-Yokuts

	Underlying	/maːt/	
⇨	Epenthesis	maːti	To repair trimoraic syllable.
	Palatalization	maːči	*ti* → *či* generally.
⇨	Apocope	maːč	Final vowels delete
	Shortening	mač	To repair trimoraic syllable.

This DY case reflects an unattested and presumably impossible type of rule interaction. Yet, as Kiparsky (e-mail communication, 7 July 1998) points out and as I will now show, quasi-Yokuts is analyzeable in sympathy theory, if information is transmitted from the sympathetic candidate to the rest of the candidate set by inter-candidate faithfulness constraints.

The basic phonology of quasi-Yokuts is given by the rankings in (36):

(36) Constraint Rankings for Quasi-Yokuts

a.	$*[\mu\mu\mu]_\sigma$ >> DEP-V	Trimoraic syllables are repairable by epenthesis.[20]
b.	$*[\mu\mu\mu]_\sigma$ >> MAX-μ	Trimoraic syllables are repairable by shortening.
c.	DEP-V >> MAX-μ	Shortening is preferred to epenthesis.
d.	**ti* >> IDENT(high)	There is palatalization.

Under the assumption that codas contribute to weight, CVVC syllables run afoul of $*[\mu\mu\mu]_\sigma$. This constraint is able in principle to compel both epenthesis and shortening; which one actually occurs is determined by the ranking in (36c), which favors shortening over epenthesis. The last ranking, by deploying the ad hoc constraint **ti* above IDENT(high), accounts for the palatalization process.

To simulate the feeding DY pattern, the sympathetic candidate must be ❀*maːči*, like the intermediate stage of the serial derivation. This sympathetic candidate is chosen if the selector constraint is ✭MAX-μ, favoring the most harmonic candidate that does not show the effects of vowel shortening. And to transmit the effects of palatalization from the sympathetic candidate to the actual output form, we can call on the sympathy constraint ❀IDENT(high). By dominating its input-output counterpart IDENT(high), the sympathy constraint ❀IDENT(high) ensures that palatalization in the sympathetic candidate is repeated in the actual output form, even if not present in the input.

The following tableau confirms the details of the analysis:

(37) Quasi-Yokuts in Sympathy Theory with Inter-Candidate Faithfulness Constraints

	/maːt/	$*[\mu\mu\mu]_\sigma$	**ti*	❀IDENT(high)	IDENT(high)	DEP-V	✭MAX-μ
Transparent	a. ☜ mat			*!			*
Sympathetic	b. ❀ maːči				*	*!	✔
Opaque	c. ☞ mač				*		*
	d. maːt	*!		*!			✔
	e. maːti		*!	*!		*	✔
	f. maːč	*!			*		✔

The actual output form is *mač*. Its transparent competitor **mat* lacks the sympathetic effect of palatalization, and it is not optimal, because of high-ranking ❀IDENT(high). The sympathetic candidate ❀*maːči* is chosen for its obedience to the selector ✭MAX-μ; of all the candidates that obey the selector, it is the most harmonic, since it contains no trimoraic syllables and has palatalization before *i*. Other candidates incur fatal violations of undominated markedness constraints, so they require no further attention.

This analysis pretty effectively simulates a feeding DY derivation. The input /maːt/ is mapped onto the output *mač* through sympathetic attraction to ❀*maːči*. On the assumption that such cases are not merely non-existent but actually impossible, we have to conclude that the original faithfulness-based theory of sympathy is too powerful.

What is the source of this problem? The theory's excessive richness comes from the existence of inter-candidate faithfulness constraints like ❀IDENT(high). These constraints allow essentially any information about the sympathetic candidate to be transmitted to the actual output form. Palatalization in quasi-Yokuts is a mere side effect of a spurious epenthesis process, yet sympathetic faithfulness constraints have no difficulty in transmitting the result of palatalization from the sympathetic candidate to the actual output form. I therefore reject the whole notion of inter-candidate faithfulness constraints and here propose a more restrictive alternative.

5.2 The solution

As the earlier discussion of ❀SYM emphasized, the point of sympathy theory is to require some sort of resemblance between the output form and the sympathetic candidate. The flawed approach based on inter-candidate faithfulness involves checking this resemblance directly, using specific constraints on candidate-to-candidate correspondence. The alternative I intend to explore here compares candidates indirectly, in terms of the unfaithful input→output mappings that created them.[21] If a candidate C has a superset of the sympathetic candidate ❀C's unfaithful mappings, then C and ❀C stand in a relation of *cumulativity*: C accumulates all of ❀C's unfaithful mappings, and may add some more of its own. DY derivations, including quasi-Yokuts, are non-cumulative — *mač* does not have a superset of ❀*maːči*'s unfaithful mappings.

To implement this idea formally, we require a definition of what an "unfaithful mapping" is, and we need a metric for comparing the sets of unfaithful mappings incurred by two candidates (one of which is the sympathetic candidate) derived from the same input. Each of these prerequisites will be addressed in turn.

Unfaithful mappings are a tokenized version of faithfulness, specifying the type and locus of unfaithfulness more precisely than constraints do. In some cases, faithfulness constraints may disregard differences in type of unfaithfulness. For example, the epenthetic mappings from /ai/ to *a.ʔi* versus *a.ti* are distinct, but both simply incur a violation of DEP. And except for certain prominent positions (Beckman 1997, 1998), faithfulness constraints are indifferent to the locus of violation. For example, the same type of faithfulness violation — deletion of a segment, a violation of MAX — is involved in mapping /pap/ to *pa* or *ap*, but the loci of violation are different.

Unlike faithfulness per se, a fully characterized unfaithful mapping specifies exactly how input and output differ, resolving all potential ambiguities. In the case of constraints like DEP, the resolution is obvious: DEP(*ʔ*) and DEP(*t*) are two distinct unfaithful mappings, but they are not distinct constraints. To distinguish the locus of each unfaithful mapping, I will index elements of the input. We can therefore talk about two distinct unfaithful mappings affecting /$p_1a_2p_3$/: MAX@1, which yields *ap*, and MAX@3, which yields *pa*. In this way, the locus of faithfulness violation is always relativized to the input, and thus it is commensurable across candidates.[22]

The locus of epenthesis is usually defined on the output; to keep things simple, it would be convenient to have a way of talking about the locus of epenthesis relative to the input. Assume that the input XY is equivalent to XεY, where "ε" is the null character. An output epenthetic segment stands in correspondence with an input ε, with one or more ε's supplied as needed for epenthetic correspondence to a set of input-equivalents. The ε symbols will be indexed relative to the segment on their left, if any: for example, input /a/ is equivalent to /$\varepsilon_{0\text{-}1}a_1\varepsilon_{1\text{-}1}\varepsilon_{1\text{-}2}$/ which underlies the output candidate $ʔ_{0\text{-}1}a_1ʔ_{1\text{-}1}i_{1\text{-}2}$.

Any output candidate from a given input is almost fully characterized by the set of unfaithful mappings that yield it. “Almost” fully characterized, because candidates can differ in properties that are not governed by faithfulness and thus do not involve unfaithful mappings. The most obvious such property is syllabification, discussed below in section 6. Apparent from this, though, GEN could be thought of as emitting various sets of unfaithful mappings *qua* candidates. Distinct candidates will be associated with distinct sets of unfaithful mappings, and these sets provide the basis for a metric of similarity between candidates — a metric that can replace inter-candidate faithfulness constraints

Let U_{Cand} stand for the set of unfaithful mappings that relate some input to the output candidate *Cand*. We are interested in comparing the sets of unfaithful mappings U_{Cand1} and U_{Cand2} associated with the candidates Cand1 and Cand2, respectively. There are four situations to consider:

- $U_{Cand1} = U_{Cand2}$. In this case, Cand1=Cand2 (except for properties like syllabification, as noted). Each is trivially *cumulative* with respect to the other.
- $U_{Cand1} \subsetneq U_{Cand2}$. In this case, Cand1 and Cand2 are different but comparable.[23] Cand2 is non-trivially cumulative with respect to Cand1; that is, Cand2 accumulates Cand1's unfaithful mappings, and adds some more of its own.
- $U_{Cand1} \supsetneq U_{Cand2}$. Cand1 and Cand2 are likewise comparable, and, symmetrically, Cand1 is cumulative with respect to Cand2.
- $U_{Cand1} \not\subset U_{Cand2}$ and $U_{Cand1} \not\supset U_{Cand2}$. Then Cand1 and Cand2 are non-comparable, and there is no relationship of cumulativity between them.

In short, cumulativity is defined in terms of a subset relation over unfaithful mappings.

The theory of partial orderings provides a more perspicuous way of looking at these inter-candidate relations. The candidates that is most faithfully mapped — identical to the input — stands at the top of the partial ordering, and below it is a rank of candidates each of which has a single unfaithful mapping. Below that is a rank of candidates each of which combines two of the unfaithful mappings from the first row, and so on. Partial orderings are best seen diagrammatically, as in the following fragment of the Hebrew *deše* example from McCarthy (1999b). Underlying /dešʔ/ is mapped onto surface *deše* by two unfaithful mappings, epenthesis of *e* and and deletion of *ʔ*. Interesting candidates include sympathetic ❀*dešeʔ* and the transparent competitor *deš*, where *ʔ* has deleted without the seemingly superfluous epenthesis process.

(38) Partial Ordering Diagram for Hebrew /$\text{d}_1\text{e}_2\text{š}_3\text{ʔ}_4$/

	dešʔ			*Faithful*	
e		i	e		i
☜deš		❀dešeʔ	MAX@4		DEP@3-1
i		e	i		e
	☞deše			*Both*	

Obviously, this is just a tiny portion of the candidate set. Standing at the top of the partial ordering is the most faithful candidate (see Moreton (1996/1999) on why such a candidate must exist). Below it, on the first tier, are candidates with a single unfaithful mapping, including deletion of /ʔ/ or epenthesis of /e/. At the next level down is the candidate *deše*, the actual output, which has suffered both of these unfaithful mappings.

The candidate standing at the top, *dešʔ*, is comparable with all other candidates, and all other candidates accumulate its unfaithful mappings. (That is because the fully faithful candidate has no unfaithful mappings, and every set is a superset of the null set.) The actual output *deše* is comparable with all of the candidates shown (though not with all possible candidates), and so it is cumulative with respect to the sympathetic candidate ❀*dešeʔ*. Significantly, *deše*'s transparent competitor **deš* is not cumulative with respect to the sympathetic candidate ❀*dešeʔ*. This non-cumulativity proves to be fatal.

A similar diagram can be constructed for another of the examples in McCarthy (1999b), the failure of open-syllable vowel raising to occur in Bedouin Arabic words like *ba.du*, derived by glide vocalization from underlying /badw/:

(39) Partial Ordering Diagram for Bedouin Arabic /$b_1a_2d_3w_4$/

	❀badw.			*Faithful*	
	e i			e i	
bidw.		☞ba.du	IDENT(high)@2		DEP-μ@4
	i e			i e	
	☜bi.du			*Both*	

The sympathetic candidate is one in which the underlying glide has not vocalized, and so raising, which occurs in an open syllable, is not motivated. In this case, then, the sympathetic candidate is faithfully mapped from the input. The actual output form has glide vocalization but not raising; it competes with **bi.du*, which transparently has both.

As usual, the faithfully-mapped candidate ❀*badw* is comparable with all other candidates, and moreover all other candidates vacuously accumulate its empty set of unfaithful mappings. The difference between the actual output *badu* and its transparent competitor **bidu* is that *badu* is closer, in terms of shared unfaithful mappings, to the sympathetic candidate.

These examples give a pretty good idea of how the revised sympathy system will work: only candidates that accumulate the unfaithful mappings of the sympathetic candidate are in sympathy with it, and among those candidates it is best to be closest, in terms of shared unfaithful mappings, to the sympathetic candidate. There are various ways to implement this system formally, and here I will take an approach suggested to me by Alan Prince. Replacing the diverse inter-candidate correspondence constraints in the earlier theory, there are just two sympathy constraints (per selector). They compare a candidate's accumulated unfaithful mappings to those of the sympathetic candidate:

(40) Cumulativity[24]

Given a sympathetic candidate ❀-Cand, to evaluate a candidate E-Cand:

a. ❀CUMUL

E-Cand is cumulative with respect to ❀-Cand. That is, $U_{❀\text{-Cand}} \subseteq U_{\text{E-Cand}}$.

b. ❀DIFF

Every unfaithful mapping incurred by E-Cand is also incurred by ❀-Cand.

That is, assign one violation mark for every member of the set $U_{\text{E-Cand}} \backslash U_{❀\text{-Cand}}$.[25]

c. Fixed Universal Ranking

❀CUMUL >> ❀DIFF

❀CUMUL evaluates each candidate categorically for whether it accumulates all of the sympathetic candidate's faithfulness violations. (See the Appendix for a refinement of this definition.) ❀DIFF evaluates candidates gradiently for how far they are from the sympathetic candidate in terms of unshared faithfulness violations. The fixed ranking places the more stringent test universally higher. Because of this fixed ranking, evaluation by ❀DIFF will only be relevant when ❀CUMUL is not decisive, thereby ensuring that only comparable candidates (in the technical sense used above) are actually compared by ❀DIFF.

Here is a shortcut that uses a diagram like (38) or (39). If there is a purely downward path from ❀-Cand to E-Cand, then E-Cand satisfies ❀CUMUL and the number of links in that path is the number of marks on ❀DIFF that E-Cand incurs. If there is no purely downward path from ❀-Cand to E-Cand, then E-Cand violates ❀CUMUL, and so its performance on ❀DIFF is of no consequence. In (38), ❀CUMUL correctly favors *deše* over its transparent competitor **deš*, relative to the sympathetic candidate ❀*dešeʔ*. And in (39), both *ba.du* and **bi.du* are cumulative with respect to the sympathetic candidate ❀*badw.*, but *ba.du* is closer, in terms of shared unfaithful mappings, so *ba.du* receives one "*" from to ❀DIFF to **bi.du*'s two "*"s.

Intuitively, these two notions, cumulativity and distance in terms of shared faithfulness violations, are analogous to criteria that have sometimes been imposed on serial derivations. The requirement that derivations be monotonic (as in Declarative Phonology (see Scobbie 1993, and references there), meaning that they take a steady path away from the input, never back-tracking, is roughly equivalent to saying that later steps of the derivation are cumulative, in the sense described above, with respect to earlier steps. And derivational economy, meaning that the length of the derivational path is minimized (Chomsky 1995: 138ff.), approximates the effect of checking the number of unshared faithfulness violations. The difference, of course, is that these notions have not previously been couched in terms of faithfulness, which is unique to OT.

Back to Bedouin Arabic. Recall that sympathy must favor *(k.tí).bat* over transparent **(kát).bat* relative to the sympathetic candidate ❀*(ki.tí).bat*. The definition of ❀CUMUL in (40) does exactly that. Consider the sets of unfaithful mappings associated with these three candidates:[26]

(41) Unfaithful Mappings Relative to Input /$k_1a_2t_3a_4b_5a_6t_7$/

	Candidate	$U_{Candidate}$		
☞	(k.tí).bat	{ STAY-ADJ(Dur)@2,	STAY(Dur)@2,	STAY(Dur)@4 }
❀	(ki.tí).bat	{	STAY(Dur)@2,	STAY(Dur)@4 }
☜	(kát).bat	{ STAY-ADJ(Dur)@4,		STAY(Dur)@4 }

The desired output form *(k.tí).bat* has a proper superset of the sympathetic candidate's unfaithful mappings, so it obeys ❀CUMUL. The transparent competitor **(kát).bat* has a partly disjoint set of unfaithful mappings from the ❀-candidate's. They are non-comparable or, equivalently, non-cumulative, and so **(kát).bat* violates ❀CUMUL. The following tableau updates (35) to reflect these developments:

(42) ❀CUMUL Applied to /katab-at/ → *(k.tí).bat*, **(kát).bat*

	/katab-at/	GRPHARM	❀CUMUL	✯STAY-ADJ(Dur)	REDUCE	STAY(Dur)
Opaque	a. ☞ (k.tí).bat			*	*	**
Sympathetic	b.❀ (ki.tí).bat	*!		✓	**	**
Transparent	c. ☜ (kát).bat		* !	*		*

Because ❀CUMUL is decisive, performance on ❀DIFF is irrelevant, and so the latter constraint is not shown in the tableau.

This is a good point to summarize the discussion. The original implementation of sympathy theory posited a set of inter-candidate faithfulness constraints which permit any property (so long as it can be named in a correspondence constraint) to be transmitted from the sympathetic candidate to the actual output form. Here I have proposed a more restrictive alternative, in which the only information that can be transmitted from the sympathetic candidate is the set of its unfaithful mappings.

The central role of cumulativity in this revised theory of sympathy is the key to explaining the impossibility of DY derivations, particularly the quasi-Yokuts case in (37). DY serial derivations are, by their very nature, non-cumulative; rather than monotonically increasing the unfaithful mappings relative to the input, they proceed non-monotonically, introducing an unfaithful mapping at one stage and then undoing it at a later stage, as in (16). Non-cumulativity makes a simulation in terms of the revised sympathy theory impossible — a welcome result, since the need for DY derivations is not supported empirically, as I argued in sections 2 and 3 above.

To see this concretely, consider the following diagram, which organizes the quasi-Yokuts candidates in (37) according to their unfaithful mappings:

(43) Partial Ordering Diagram for Quasi-Yokuts /$m_1a_2a_3t_4$/

maːt	*Faithful*
q | p	q | p
maːti maːč ☜mat	DEP-V@4-1 IDENT(high)@4 MAX-μ@3
y t y t	y t y t
❀maːči ☞mač	*Both* *Both*

It is immediately evident that the intended output *mač* does not accumulate the unfaithful mappings of the sympathetic candidate ❀*maːči*. Top-ranked ❀CUMUL extinguishes all non-cumulative candidates, leaving only ❀*maːči* as a viable candidate. Thus, ❀CUMUL is equally fatal to *mač* and its transparent competitor **mat*. This means that the quasi-Yokuts DY derivation cannot be simulated under the revised sympathy theory.

The quasi-Yokuts example highlights a general result. Under the revised theory of sympathy, the relation between the output and the sympathetic candidate is one of cumulative unfaithful mappings. DY derivations, whether implemented serially or simulated with an underlying-sympathetic-surface triplet, are inherently non-cumulative. The revisions to sympathy theory have made it more restrictive, by limiting the kinds of information that can be extracted from the sympathetic candidate. One indication of this greater restrictiveness is the impossibility of reproducing DY derivations; others no doubt remain to be discovered.

6 Cumulativity, faithfulness, and syllable structure

6.1 Overview of the issue

Cumulativity is defined in terms of shared unfaithful mappings. In fact, the candidates themselves can be described in terms of the unfaithful mappings that produced them — up to a point. Candidates may also differ in properties that are phonologically relevant but not governed by faithfulness constraints. Here's the difference. Any phonological property that is independently contrastive in the phonology of some language must be protected by faithfulness constraints provided by UG, and each breach of a faithfulness constraint will count as an unfaithful mapping for the purposes of assessing cumulativity. But properties that are never contrastive in the phonology of any language are not subject to faithfulness constraints. For example, Keer (1999) argues, from the observation that tautomorphemic true and fake geminates are never contrastive (Hayes 1986, McCarthy 1986), that fusion of two adjacent identical segments exacts no cost in faithfulness. Any such faithfulness-free mappings will be irrelevant to determining how well a candidate performs on ❀CUMUL and ❀DIFF.

It is virtually a truism that syllabification is never contrastive in any language. (This claim has some subtleties, to be discussed later, involving juncture effects and distinctions of quantity or syllabicity.) No language is known to contrast tautomorphemic *pa.ta* with *pat.a* or *pa.kla* with *pak.la* (though see van Oostendorp (this volume) for a possible counterexample.) This observation is usually taken to mean that syllabification is absent from underlying representations (e.g., Blevins 1995: 221, Clements 1986b: 318, Hayes 1989: 260). It is, however, more in keeping with OT's thesis of richness of the base (Prince and Smolensky 1993) to assume that underlying representations may be syllabified or not and in diverse ways — freely but pointlessly, since no constraints of UG lobby for the conservation of underlying syllabification:

(44) Faithfulness-Free Syllabification

No constraints of UG demand faithfulness to syllables *per se*.

This section explores the implications of (44) for cumulativity and opacity.

A derivation is cumulative if it monotonically increases its unfaithful mappings. Cumulative derivations are in general permitted, but non-cumulative derivations are not, for the reasons given in section 5. If the thesis of faithfulness-free syllabification is correct, then syllabification is irrelevant to cumulativity, and so it should be possible to find real derivations where syllabification changes, non-monotonically. These derivations will have something of the look and feel of the unattested feeding DY derivations, but they will involve non-monotonicity *only* in faithfulness-free syllabification.[27]

The example back in (3) is a hypothetical, though undoubtedly authentic-appearing, instance of this. A genuine case comes from Clements's (1986a) analysis of quantity in Luganda (cf. Rosenthall 1994, Wiltshire 1992). In Luganda, vowels are always long before prenasalized consonants: *ku-liːnda* 'to wait', *mu-leːnzi* 'boy', *muː-ntu* 'person', *baː-ntu* 'people'. Clements argues that this is an effect of compensatory lengthening: the nasal is first syllabified as a weight-bearing coda, and then is joined to the following consonant, leaving a stray weight-unit to be filled by spreading from the preceding vowel. Here is a derivation, substituting moras for the CV weight units that Clements uses:

(45) Luganda Derivation with Resyllabification

```
                                  μ   μ
                                  |   |
⇨     Syllabification           mun.tu

                                  μμ  μ
                                  ||  |
      Weight by Position        mun.tu

                                  μμ  μ
                                  |   |
⇨     Prenasalization           mu .^ntu

                                  μμ  μ
                                  |/  |
      Spreading                 mu .^ntu
```

Just like the feeding DY cases, the intermediate stage plays a crucial role, since it supplies the context for Weight by Position (Hayes 1989), which assigns a mora to the nasal.

This derivation is cumulative, even though it has the DY-like step of Prenasalization undoing the effects of earlier Syllabification. To show that formally, it is necessary to sketch a partial account of Luganda within OT, using the revised theory of sympathy to deal with the opaque interaction between the process assigning positional weight and the process creating prenasalized consonants.[28]

There are no (non-geminate) codas in Luganda. Potential nasal codas, like the *n* in /muntu/, are disposed of by coalescence with a following consonant, in violation of the faithfulness constraint UNIFORMITY (McCarthy and Prince 1995) (but cf. Keer 1999). These observations motivate the following ranking:

(46) NO-CODA, MAX >> UNIFORMITY

/muntu/	NO-CODA	MAX	UNIF
a. ☞ muː.ⁿtu			*
b. mun.tu	* !		
c. muː.tu		* !	

Rankings like this, where a markedness constraint and MAX together dominate UNIFORMITY, are typical of coalescence phenomena (Causley 1997, Gnanadesikan 1995, 1997, Lamontagne and Rice 1995, McCarthy to appear, McCarthy and Prince 1995, Pater 1999).

This analysis is not sufficient, however, because there is an element of opacity in Luganda coalescence, as I have noted. The mapping /muntu/ → *muː.ⁿtu* involves a seemingly gratuitous violation of DEP-μ, a violation that the transparent output form **mu.ⁿtu* would have avoided. This is a sympathy effect, induced by the sympathetic candidate ❀*muN.tu* (where capitalization marks the *N* as moraic, as in the second step of (45)). The selector constraint is ✯UNIFORMITY, which is obeyed by ❀*muN.tu* and violated by the actual output form *muː.ⁿtu*. And to ensure that the sympathetic form is the most harmonic candidate among those that obey the selector, certain additional rankings among as-yet unranked constraints are necessary. One of these is MAX >> NO-CODA, so that ❀*muN.tu* is more harmonic than **mu.tu*. Another deploys weight by position (WXP) above DEP-μ, so that ❀*muN.tu*, with a moraic coda, is more harmonic than **mun.tu*, with a non-moraic coda.

This much establishes the essential background for discussing the sympathy effect itself. Sympathy must favor opaque *muː.^{n}tu* over transparent **mu.^{n}tu* relative to the sympathetic candidate ❀*muN.tu*. To check whether it does, the first step is to assemble the sets of unfaithful mappings for these candidates:

(47) Unfaithful Mappings Relative to Input /$m_1u_2n_3t_4u_5$/

	Candidate	$U_{Candidate}$	
☞	muː.^{n}tu	{ DEP-μ@1-1,[29]	UNIF@3&4 }
❀	muN.tu	{ DEP-μ@1-1	}
☜	mu.^{n}tu	{	UNIF@3&4 }

The intended output form has a superset of the sympathetic candidate's unfaithful mappings; they are therefore in a relationship of cumulativity. But the transparent competitor does not accumulate the sympathetic candidate's unfaithful mappings. Therefore, ❀CUMUL will favor *muː.^{n}tu* over **mu.^{n}tu*, exactly as desired. The following tableau completes the argument at the level of formal detail:

(48) Luganda /muntu/ → *muː.^{n}tu*

/muntu/	MAX	WXP	NO-CODA	❀CUMUL	✭UNIF	DEP-μ
a. ☞ muː.^{n}tu					*	*
b. ❀ muN.tu			* !		✓	*
c. ☜ mu.^{n}tu				* !	*	
d. muː.tu	* !				✓	*
e. mu.tu	* !				✓	
f. mun.tu		* !	*	*	✓	

Several candidates obey the selector constraint, ✭UNIFORMITY; of those, ❀*muN.tu* is most harmonic, so it is chosen as the sympathetic candidate. Through the constraint ❀CUMUL, the sympathetic candidate bestows its favor on *muː.^{n}tu*, which accumulates its unfaithful mappings, over **mu.^{n}tu*, which does not.[30]

This analysis succeeds under the assumptions that (i) cumulativity is defined in terms of shared unfaithful mappings and (ii) syllabification is not an unfaithful mapping — i.e., (44). If syllabification were to be counted as an unfaithful mapping, then the the record of unfaithful mappings for each candidate would have to be augmented as follows:

(49) Unfaithful Mappings Relative to /$m_1u_2n_3t_4u_5$/ under Wrong Assumption about Syllabification

	Candidate	$U_{Candidate}$
☞	muː.ⁿtu	{ Ons@1, DEP-μ@1-1, Nuc@2, UNIF@3&4, Ons@3, Ons@4, Nuc@5 }
❀	muN.tu	{ Ons@1, DEP-μ@1-1, Nuc@2, Cod@3, Ons@4, Nuc@5 }
☜	mu.ⁿtu	{ Ons@1, Nuc@2, UNIF@3&4, Ons@3, Ons@4, Nuc@5 }

Observe that the sympathetic candidate and the intended output differ on the syllabification of /n/: Cod@3 vs. Ons@3. Thus, there is no cumulativity relation between these candidates, if syllabification is reckoned in the determination of cumulativity. This means that the intended output and its transparent competitor **mu.ⁿtu* both violate ❀CUMUL. This tie is disastrous; since they also tie on ❀DIFF (each incurs two marks), the decision falls to low-ranking DEP-μ (see (48)), which *muː.ⁿtu* fatally violates. That is the wrong result.

This argument shows why, as a matter of descriptive necessity, syllabification cannot be reckoned as an unfaithful mapping. What remains is to fill in the details, hinted at earlier, of how this premise fits into phonology generally. As I noted, saying that syllabification is not governed by faithfulness constraints entails that no contrast in syllabification can be preserved in the mapping from underlying to surface representations. The main challenges to non-contrastive syllabification are these:

> *Grammatically-conditioned contrast*. Morphemic juncture can produce syllabificational contrasts, as in well-known examples like *lightning/lightening* or *nitrate/night-rate*.

Phonologically-derived contrast. In Barra Gaelic, CV sequences derived by epenthesis are syllabified differently from underlying CV sequences.

Contrast in quantity or syllabicity. Consonant gemination has obvious consequences for syllabification. And contrasts between glides and vowels have been reported for Berber, Ilokano, and Spanish, *inter alia* (Guerssel 1986, Harris 1987, Hayes 1989, Levin 1985, Rosenthall 1994).

When none of these conditions obtain, syllabification does appear to be reliably non-contrastive, as in the examples of tautomorphemic *pa.ta*/*pat.a* and *pa.kla*/*pak.la* that were cited earlier. I will take each of these factors in turn, briefly showing how they are compatible with the thesis in (44) that syllabification is not regulated by faithfulness constraints.

6.2 Grammatically-conditioned contrast

Grammatically-conditioned contrasts in syllabification have been extensively studied within OT. One important source of grammatically-conditioned syllabification contrast is alignment (McCarthy and Prince 1993a). Alignment constraints can require that a segment standing at the edge of a morphological constituent, such as the stem, also stand at the edge of a prosodic constituent, such as the syllable. In English, for example, ALIGN-LEFT dominates NO-CODA, to ensure that the stem-initial *r* of *rate* is also word- and syllable-initial in *night-rate*. Where ALIGN-LEFT is irrelevant, though, as in tautomorphemic *nitrate*, the ranking of NO-CODA above *COMPLEX-ONSET will force onset maximization. A surface syllabification contrast is the result, but it does not require constraints demanding faithfulness to syllabification.

Output-output faithfulness constraints need to be considered as another potential source of grammatically-conditioned contrasts in syllabification. A central thesis of

Transderivational Correspondence Theory (Benua 1997b) is that output-output faithfulness constraints have the same formal properties as input-output (or base-reduplicant) faithfulness constraints. So if there are no constraints enforcing faithfulness to syllabification in input→output mappings, then there can be no such constraints on output→output mappings either.

English phonology is a good place to look for potential counterexamples to this thesis. The challenge comes from syllabic "closure" cases like *lightning*/*lightening* or *siren*/*siring*, where the sonorant is syllabic only before a Level II suffix (see, among others, Benua 1997b, Borowsky 1993, Harris 1990, Mohanan 1985). But even in these cases, it does not seem to be necessary to invoke faithfulness constraints on syllabification *per se*. Alignment constraints are one possible line of attack; another is moraic faithfulness. Below, I argue that faithfulness to moras, rather than syllables, is the basis of contrasts in syllabicity. The syllabic *n* of *lighten* bears a mora, under one view of syllabification. Faithfulness to this mora in the output-output dimension (i.e., OO-MAX-μ) will ensure its preservation in the derived form *lightening*. Moraic faithfulness is here a partial surrogate for syllabic faithfulness, and surrogacy appears to be enough for known cases. Of course, this surrogate also opens the possibility of introducing illegitimate syllabic contrasts through the moraic back-door; that issue is also discussed below.

A final grammatical circumstance that is relevant to syllabic faithfulness is reduplication. Reduplication never copies syllables (Marantz 1982, McCarthy and Prince 1986, 1990a, Moravcsik 1978). That is to say, no language has a single reduplicative morpheme which copies the initial *ta* of *ta.pi* and the initial *tak* of *tak.pi*. A necessary (but not sufficient) condition for excluding this possibility is that UG contain no constraints enforcing faithfulness to syllables on the base-reduplicant dimension. Again, this correlation between base-reduplicant faithfulness and input-output faithfulness is expected under correspondence theory (McCarthy and Prince 1995, 1999).

6.3 Phonologically-conditioned contrast

In Barra Gaelic, the sequence $\breve{V}_1CV_2$ is syllabified differently depending on the provenance of V_2 (Beckman 1998, Borgstrøm 1937, 1940, Bosch 1991, Bosch and de Jong 1998, Bosch and de Jong 1997, Clements 1986b, Green 1997: 230-1, Kenstowicz and Kisseberth 1979): if V_2 is underlying then the syllable boundary falls after C, but if V_2 is epenthetic then the syllable boundary falls before C. As a consequence there are surface near-minimal pairs differing in syllabification, like *ar.an* 'bread' (from /aran/) and *a.ram* 'army' (from /arm/).

Following Clements, Blevins sketches a plausible derivational analysis:

(50) Derivational Analysis of Barra Gaelic (after Blevins 1995, 231)

Underlying	/aran/	/arm/
Syllabification	a.ran	a.rm
Attraction	šæR.ak	—
Epenthesis	—	a.ram

The source of the surface contrast is the counter-feeding order between Attraction and Epenthesis. Attraction makes the stressed initial syllable heavy by drawing in the following consonant as a moraic coda. Because Epenthesis applies later, the onset of the epenthetic syllable cannot be attracted away. A surface contrast in syllabification is the result.

Clearly, Barra Gaelic does not depend on syllabic faithfulness in the input→output mapping, so it presents no difficulties for my main premise. And in any case, there is good reason to doubt that the story in (50) is correct and complete. New phonetic evidence developed by Bosch and de Jong (1998, 1997) shows that epenthesis leads to a difference in stress: *ár.an* vs. *a.rám*. It may be that this difference in stress is directly responsible for the syllabification difference (see also Beckman 1998).

6.4 Contrasts in quantity and syllabicity

It is widely though not universally accepted that contrasts of quantity and syllabicity are represented by deploying moras in underlying representation (see Davis this volume and Hayes 1989, McCarthy and Prince 1988, Rosenthall 1994, Sherer 1994). Faithfulness to underlying moras, thanks to constraints like DEP-μ and MAX-μ, ensures that these underlying distinctions are maintained faithfully at the surface. Indeed, the analysis of Luganda above shows that insertion of a mora does constitute an unfaithful mapping, a result that is consistent with the role of moras in representing contrasts.

To complete the picture, though, it is necessary to show that faithfulness to underlying moras does not offer a back-door into the non-occurring *pa.ta*/*pat.a* or *pa.kla*/*pak.la* contrasts. This is not an easy undertaking: in its most general form, the claim is that no arrangement of underlying moras on the tautomorphemic string /pata/ will map onto a surface *pa.ta*/*pat.a* distinction under any permutation of the constraints of UG (and likewise for *pa.kla*/*pak.la*). Rather than solve this problem in its most general form, I propose here to address a modest-sized piece of it, the impossibility of having a language with the mappings in (51):

(51) Moraic Faithfulness as Surrogate for Syllabic Faithfulness

a. /pata/ → *pa.ta*

b. /paTa/ → *paT.a*

(Recall that capital *T* stands for *t* associated with a mora.) A language with mappings like these would be one in which faithfulness to moras in underlying representation produces contrast in syllables in surface representation. To account for the impossibility of tautomorphemic syllabic contrast, it is necessary (though not sufficient) to universally rule out the system with these mappings.

In OT, with richness of the base, the way to rule out a mapping is to find a more harmonic mapping. The way to rule out a mapping universally is to make sure that there is *always* a more harmonic mapping, under *any* permutation of the constraints of UG. Suppose that UG consists of only the following constraints:

(52) A Limited UG Constraint Set

a. ONSET

$*[_\sigma V$

b. NO-CODA

*C.

c. $*\mu_{Cons}$ (Sherer 1994: 26)

$*[C]_\mu$ (Consonants may not be parsed as moraic.)

d. WxP

If C_i., then $[C_i]_\mu$ (Coda consonants must be parsed as moraic.)

e. Faithfulness Constraints

MAX

MAX-μ

DEP-μ

This constraint set will map input /pata/ onto output *pa.ta* under any ranking (cf. Prince and Smolensky 1993). But it will also map input /paTa/ onto geminate *paT.ta* or simplex *pa.ta*, depending on the disposition of MAX-μ relative to the structural constraints. No permutation will produce *paT.a* or *pat.a* from input /paTa/, because there is no antagonistic constraint to offset their violations of ONSET. Therefore, the illicit syllable-structure contrast is not obtainable from these inputs under this theory of UG.[31]

Other theories of UG may have other consequences. For instance, if UG contains a constraint that specifically militates against geminate or ambisyllabic consonants (as in

Rosenthall (1994) or Beckman (1998)), then the illicit contrast is easily obtained simply by ranking the anti-geminate constraint above ONSET. But the fact that some phonemic inventories lack geminates does not necessarily mean that UG has a constraint against geminates specifically. The theory in (52) can rule out geminates indirectly, under the following ranking permutations:[32]

(53) Some Geminate-less Permutations of (52)

a. No Geminates, No Codas Whatsoever

All others >> MAX, MAX-μ

b. Moraic Codas, but No Geminates

ONSET, WXP, MAX >> *μ_{Cons}, DEP-μ, NO-CODA >> MAX-μ

c. Non-moraic Codas and No Geminates

ONSET, MAX, *μ_{Cons}, DEP-μ, >> MAX-μ, NO-CODA, WXP

The source of geminatelessness under (52) is the ranking *μ_{Cons} >> MAX-μ, and this ranking will never aid and abet the illicit mapping /paTa/ → *paT.a*. For further discussion, see Keer (1999: 48ff.).

6.5 Summary

To summarize the results of this section, I have argued that there are no constraints enforcing faithfulness to syllables per se. A theory-internal argument, based on applying revised sympathy to cases like Luganda, was supported by theory-external observations about non-occurring contrasts and impossible reduplicative patterns. Several challenges to this thesis were also addressed: grammatically-conditioned and phonologically-derived contrasts, and distinctions of quantity and syllabicity. Finally, I showed that moraic faithfulness, which is necessary to maintain contrasts in quantity and syllabicity, need not lead to illicit syllabification contrasts.

7 Conclusion

The serial derivation, although it is a central concept of generative phonology, has been little studied. A rare exception is Pullum's (1976) work on the Duke-of-York gambit, a type of derivation where the output returns to the same place as the input. Though serial rule-based phonology predicts the existence of DY derivations, they do not seem to occur, except as descriptive artifacts of serialism's commitment to rule prioritization through ordering.

The principal goal of this article has been to explain the impossibility of DY derivations in their most general form. The argument here is embedded within OT, and more specifically within the extension to OT called sympathy, which addresses opaque interactions among processes. The key is a revision of sympathy theory, changing the means by which information is transmitted from the sympathetic candidate to the output form. Instead of inter-candidate faithfulness constraints, I have argued for a considerably more restrictive hypothesis: candidates are compared for their faithfulness violations. The actual output must accumulate the faithfulness violations of the sympathetic candidate. This notion of cumulativity is what separates real derivations from non-existent DY derivations.

The article concluded with some study of the role of syllabification in derivations. Syllabification, I argued, is not governed by faithfulness, and so it does not figure in the reckoning of cumulativity. Theoretical and empirical consequences of this view were presented.

The results presented here suggest that familiar notions like the serial derivation, which might seem to have little or nothing left to offer, bear close re-examination. It is perhaps significant that the questions raised by Pullum have not been much studied in the intervening decades; it is certainly significant that these questions still claim our attention.

Appendix: Sympathy, Cumulativity, and Harmonic Ascent[33]

Classic OT grammars share a property of *harmonic ascent* (Moreton 1996/1999). A classic OT grammar, following Prince and Smolensky (1993), is a ranking of markedness and faithfulness constraints, and nothing else. Because violation is minimal, unfaithfulness is only possible to achieve markedness improvement relative to some language-particular ranking of the markedness constraints in UG. So, if a language has an unfaithful mapping /A/ → B, then B must be less marked, relative to that language's hierarchy, than the fully faithful candidate A. (See Moreton 1996/1999 for a formal proof of this result and discussion of its empirical consequences, such as the impossibility of circular chain-shifts.)

Elliott Moreton (personal communication, 5 May 1998) has shown that the harmonic ascent property also holds of classic OT grammars to which a single sympathy constraint has been added. To see why, assume that this claim is not true — that is, assume that input /A/ maps onto output B even though B is more marked than (output) A. *Ex hypothesi*, this mapping occurs by virtue of sympathy to some third candidate C which is selected by the faithfulness constraint ✯F. The following tableau shows the imagined situation, where A is less marked than B, yet B is the output. Observe too that C must be even less marked than A; since C and A both obey the selector ✯F, C must be less marked than A if it is to be the **most harmonic** candidate that obeys the selector:

(54) Harmonic Descent with One Sympathy Constraint (partial tableau)

	/A/	M	✯F
Faithful	a. A	*	✓
Harmonic Descent	b. ☞ B	**	*
Sympathetic	c. ❀ C		✓

Any sympathy constraint that will ensure B's victory must be ranked above M. It must also choose B over both A and C, since B is harmonically bounded by both. But no sympathy constraint can possibly do this: how could sympathetic resemblance to C somehow disfavor C itself? There is an obvious contradiction here, and so harmonic descent is impossible with a single sympathy constraint. This is a desirable result, since harmonic ascent is arguably a welcome consequence of OT.

Moreton goes on to show, however, that harmonic descent is possible if there are two sympathy constraints with two sympathetic candidates and two selectors (for a worked-out example of opacity with two selectors, see McCarthy 1999b). The trick is that each sympathy constraint is called on to exclude the other's sympathetic candidate, as in the following schematic tableau:

(55) Harmonic Descent with Two Sympathy Constraints

	/A/	❀S_{F1}	❀S_{F2}	M	✯F_1	✯F_2
Faithful	a. A	*!		*	✓	✓
Harmonic Descent	b. ☞ B			**	*	*
Sympathetic	c. ❀$_{F1}$ C		*!		✓	*
Sympathetic	d. ❀$_{F2}$ D	*!			*	✓

The sympathy constraints ❀S_{F1} and ❀S_{F2} are indexed to the faithfulness constraints that serve as their selectors. Each rules out the other's sympathetic candidate, and in addition ❀S_{F1} ensures that B is more harmonic than A, despite B's worse performance on the markedness constraint M.

The undesirable result in (55) cannot be avoided in the correspondence-based sympathy theory of McCarthy (1998), but it can be eliminated by refining the definition of cumulativity in (40). The faithfulness relationships among the various candidates are given by the following diagram (cf. (38), (39), and (43)):

(56) Partial Ordering Diagram for (55)

Under the definition (40), only B accumulates all of the unfaithful mappings incurred by both ❀C and ❀D, so only B satisfies both ❀CUMUL$_{F1}$ and ❀CUMUL$_{F2}$. But suppose (40) is modified as follows:

(57) Cumulativity (Revised)

a. ❀CUMUL

E-Cand and ❀-Cand are comparable. That is, $U_{❀\text{-Cand}} \subseteq U_{E\text{-Cand}}$ or $U_{❀\text{-Cand}} \supseteq U_{E\text{-Cand}}$.

b. ❀DIFF

Every unfaithful mapping incurred by E-Cand is also incurred by ❀-Cand, and vice-versa. That is, assign one violation mark for every member of the set $U_{E\text{-Cand}} \setminus U_{❀\text{-Cand}}$ and every member of the set $U_{❀\text{-Cand}} \setminus U_{E\text{-Cand}}$.

Revised in this way, ❀CUMUL and ❀DIFF will evaluate A and B in (56) as equally harmonic, since both stand in a super- or subset relation to the sympathetic candidates, and both are equidistant from the sympathetic candidates. This will leave the decision up to the markedness constraint M, which rules out candidate B. As a consequence of this revised definition, harmonic descent is impossible in OT with sympathy theory, just as it is in classic OT.

NOTES

* I am grateful to the participants in the Tübingen workshop, particularly the organizers, Caroline Féry and Ruben van de Vijver. I am also grateful to the UMASS weekly phonology group (Patrik Bye, Nancy Hall, Heli Harrikari, Caroline Jones, Ed Keer, Cecilia Kirk, Paul de Lacy, Meredith Landman, Ania Łubowicz, Steve Parker, and Jen Smith) and to audiences at the University of California, Santa Cruz (especially Junko Itô, Armin Mester, Jaye Padgett, and Geoff Pullum), Indiana University (especially Stuart Davis, Dan Dinnsen, and Ken de Jong), and the University of Connecticut (especially Andrea Calabrese). Special mention: Paul Kiparsky forced me to re-examine my ideas about sympathy with his example (16); Itô and Mester sent me very useful comments based on discussion in their seminar; Ed Keer, Paul de Lacy, Ania Łubowicz, Alan Prince, and Colin Wilson commented on an earlier draft; additionally, Alan Prince heard me out and offered many excellent suggestions.

1 *Bleeding* DY interactions, where the intermediate stage waits out another process, may be reducible to the vacuous type. See section 2.

2 To complete the analysis, it is also necessary to dispose of candidates where the vowel loses its rounding or the dorsal shifts to another place of articulation. I assume that such candidates are dealt with by high-ranking faithfulness constraints.

3 If the underlying form /m̉oːq/ is assumed, the Delabialization rule applies only vacuously. Vacuous application is generally regarded as sufficient to trigger the EC, however.

4 The Bedouin Arabic vowel-height alternations present the same problem: a rule raising *a* to *i* when the next syllable contains *i* must take precedence over a rule lowering *i* to *a* after a guttural consonant (see note 11). Clearly, there is no specific/general relation between these rules either.

5 According to Klokeid (1977), the word that Campbell (1973) writes as *m̉oːq* in (5c) is actually pronounced *m̉oːq*^{w}. This is unlikely to be correct — Campbell refers to his own

fieldwork with Nootka, but Klokeid evidently lacks comparable experience — but in any case it does not matter. Suppose the output is indeed *m̓oːq*^{w}. It is then necessary to supply a grammar that will map both /m̓oːq/ and /m̓oːq^{w}/ onto this output. From input /m̓oːq/, there is a DY derivation: /m̓oːq^{w}/ → *m̓oːq* → *m̓oːq*^{w}. In OT, the ranking "ROUNDING" >> "UNROUNDING" >> IDENT(round) accomplishes the same thing.

6 Even the highest-ranking constraint will be violated if Gen supplies no candidates that obey it. That situation probably never arises in phonology, where Gen meets the requirement of Inclusiveness (McCarthy and Prince 1993b: 5).

7 Merchant (1997) takes a different tack, reinterpreting the Catalan DY derivation α→β→α schema as α→β→γ, where α and γ are phonetically identical but structurally different. In his account, output *sublunar* has a single [+voice] specification shared by the *bl* cluster, but input /sublunar/ does not.

Joan Mascaró (e-mail, 24 February 2000) informs me that Harris's Catalan example *sublunar* in (12) is problematic. Except for the orthographic [b], there is evidently no reason to assume that the final consonant of the prefix *sub-* is voiced, since it also shows up as voiceless before a vowel. (The prevocalic behavior is analogous to compounds like *sud-est*.) Hence, it never occurs in a position where it could show a voicing contrast.

8 This line of attack on bleeding DY derivations would be foreclosed in any case where either UG does not supply the equivalent of the constraint *EAD or where independent evidence proves that *CA dominates *EAD. The derivation *puš* → *pɨš* → *puš* 'push' in Chomsky and Halle (1968: 294) is perhaps such a case, but the analysis is not worked out and seems dubious on its face (Zwicky 1974: 216).

9 I am grateful to Morris Halle, Harry van der Hulst, and Bill Idsardi for bringing up the Hebrew example. Compare Idsardi (1998) for an approach to these alternations based on different assumptions about the underlying representation.

10 The transcription has been simplified by suppressing indications of velar palatalization and contextual effects on the raised vowel. The data are drawn from the Ħarbi dialect carefully described and analyzed by Al-Mozainy (1976, 1981) (see also Al-Mozainy et al. 1985). (A text dictated by Al-Mozainy has been published in Ingham (1982: 112-115).) Similar data can be found in other Saudi Bedouin dialects, such as those in Abboud (1979), Johnstone (1967a, 1967b), and Prochazka (1988). The Levantine and North African Bedouin dialects described by Blanc (1970), Irshied (1984), Irshied and Kenstowicz (1984), and Mitchell (1960) differ significantly. The more familiar sedentary dialects of Arabic, such as those discussed by Kiparsky (this volume) and Wiltshire (this volume), are even more different, reflecting an ancient split between two dialect groups.

11 Al-Mozainy (1981) takes considerable care in establishing that the underlying forms given in (18) are correct. His most controversial claim is the vowel posited in the first syllable of /katab/ — underlying /a/, but always *i* or Ø on the surface. Three arguments support underlying /a/. First, a distinction between /a/ and /i/ is necessary to account for the different paradigms of /katab/ and its passive /kitib/. Observe that the presence or absence of a surface vowel between *k* and *t* is exactly complementary in these two paradigms. Second, there are paradigms where the initial *a* posited in /katab/ does show up, because there are conditions where raising is blocked (Al-Mozainy 1981, Gafos and Lombardi 1999, Irshied and Kenstowicz 1984, McCarthy 1999a):

(i) Conditions That Block Raising

Raising occurs in non-final light syllable:

/katab/ → *k**i**tab*

Except

a. Before a guttural consonant (*ʔ, h, ʕ, ħ, ʁ, χ*) or coronal sonorant (*r, l, n*)

followed by *a*:

	/saħab/	s**a**ħab	'he pulled'
	/daras/	d**a**ras	'he studied
cf.	/taʕib/	t**i**ʕib	'he got tired'
	/šarib/	š**i**rib	'he drank'

b. After a guttural consonant:

/ʔakal/	ʔ**a**kal	'he ate'

Except before high vowel:

/χasir/	χ**i**sir	'he lost'

Third, secret-language data demonstrate that the raising rule applies completely productively, with *a* or *i* in an open syllable depending on the blocking conditions just mentioned:

(ii) External Evidence for Raising and Conditions On It

a. Root-consonant-permuting secret language:

d**i**faʕ	'he pushed'	d**a**ʕaf, ʕ**a**daf, f**a**ʕad, f**i**daʕ, ...
ð**a**rab	'he beat'	ð**i**bar, r**i**bað, barað, ...
ðr**i**bat	'she beat'	rb**i**ðat, bð**a**rat, rð**i**bat, ...

b. Invented ħar-inserting secret language:

k**i**tab 'he wrote' → k**a**ħartab

12 As usual in Arabic, final superheavy syllables act like heavy penults, and so receive stress: *maktúːb* 'written', *saħábt* 'I pulled'. See Kiparsky (this volume) and Wiltshire (this

volume) for discussion.

13 Al-Mozainy (1981) reports syllabifications like *kti.bat* and *yak.tbin*, based on his own intuitions. Since surface degenerate syllables were not an option in the theory of the time, he had no choice but to assign the extra consonant to one of the visible syllables.

14 "Categorically less faithful" means that there must be a distinct high-ranking faithfulness constraint against the /A/ → C mapping. It is not enough for /A/ → C simply to accumulate more violations of the same faithfulness constraint that /A/ → B and /B/ → C violate.

15 According to Kirchner (1996), relative faithfulness is established by locally conjoining faithfulness constraints (in the sense of Smolensky 1995). Since any two faithfulness constraints can in principle be conjoined, this is an inherently richer theory than Gnanadesikan's scales.

16 See Morelli (this volume) for a fuller explanation of harmonic bounding under the rubric of "The Subset Strategy". Harmonic bounding was introduced by Samek-Lodovici (1992) and also figures prominently in Prince and Smolensky (1993: Chapter 9) and Samek-Lodovici and Prince (1999).

17 A natural idea is to attempt some sort of reformulation of GrpHarm to prefer the ΔL foot of *(k.tí).bat* over the monosyllabic H foot of **(kát).bat* (cf. Black 1991). The problem is that any such move will interfere with the syncope process. Recall that the high vowel of /samiʕ-at/ deletes to yield *(sám).ʕat*. This speculative reformulation of GrpHarm would instead favor **(s.mí).ʕat*. The following table makes this problem clear:

(i) Summary of the **(kát).bat* Problem

Input:	/katab-at/	Input:	/samiʕ-at/
Actual Output:	*(k.tí).bat*	Actual Output:	*(sám).ʕat*
Failed Candidate:	**(kát).bat*	Failed Candidate:	**(s.mí).ʕat*

The failed candidate from /katab-at/ has exactly the shape of the actual output from /samiʕ-at/, and vice-versa. This means that no markedness constraint(s) can successfully sort out these candidate-comparisons. And the faithfulness system is not helpful either — the failed candidate **(kát).bat* is in fact more faithful (because it preserves /a/) than the actual output *(k.tí).bat.*

18 Because sympathy allows a non-surface candidate to influence the outcome, it is sometimes suggested that sympathy is basically a restatement of the rule-based serial derivation. The differences between sympathy and rule-based serialism are the topic of sections 3.2 and 7 of McCarthy (1999b), and the treatment of opacity in serialized OT ("harmonic serialism") is covered in section 4 of McCarthy (2000).

19 Other approaches to opacity in OT include local constraint conjunction (Ito and Mester, this volume, Kirchner 1996) and serially ordered levels (Kiparsky, this volume, and Bermúdez-Otero 1999, Booij 1996, 1997, Clements 1997, Cohn and McCarthy 1994, Hale and Kissock 1998, Hale et al. 1998, Kenstowicz 1995, Kiparsky 1997a, 1997b, 2000, McCarthy and Prince 1993b: Appendix, Noyer 1997, Paradis 1997, Potter 1994, Roca 1997, Rubach 1997, 2000). For discussion of these and other alternatives to sympathy, see sections 6.2 and 8 of McCarthy (1999b).

20 This ranking is introduced solely to pick out the right ❀-candidate; it is otherwise unmotivated.

21 I am greatly indebted to Alan Prince for discussion of this material.

22 Though described here in segmental terms, this approach can be generalized in obvious ways to handle moras, tones, and other non-segmental structure.

23 The terms *comparable* and *non-comparable* come from the theory of partial orderings (see, e.g., Davey and Priestley 1990).

24 These constraints are relativized to the selector constraint, to allow for systems with multiple sympathetic candidates. See McCarthy (1999b) on opacity in Yokuts, for example.

25 The notation A\B (or A–B) denotes the *relative complement* of B in A. It is the set of all elements of A which do not also belong to B.

26 I have suppressed the violations of the faithfulness constraint IDENT(Stress), which are incurred when a vowel receives (or loses) a stress. Though stress is fully predictable in Bedouin Arabic, it is contrastive in some languages, and so there must be faithfulness constraints conserving it (Alderete 1998, 1999, Bye 1996, Inkelas 1999, Ito et al. 1996, McCarthy 1995, Pater 1995). Violations of IDENT(Stress), although they do not affect the outcome in in (41), are part of the package of unfaithful mappings that each candidate brings with it.

27 In fact, one can construct cases which have not only the look and feel, but even the actual form of DY derivations. Under richness of the base, underlying representations may contain syllabic structure, allowing for the possibility of derivations like /pak.la/ → *pa.kla* → *pak.la*.

28 For a comprehensive analysis of Luganda within OT, see Rosenthall (1994).

29 I assume that epenthesis on the moraic tier is treated for indexation purposes like epenthesis on the segmental tier.

30 Since **mu.^{n}tu* is literally impossible in Luganda, it is necessary to show that it cannot be attained from any input, not just /muntu/. Of particular interest is the fully faithful mapping /mu.^{n}tu/ → **mu.^{n}tu*, which must somehow be excluded.

In OT, prohibiting one mapping is a matter of ensuring that some other mapping from the same input is more harmonic. The more harmonic mapping then “occults” (Prince and Smolensky 1993) the less harmonic one. The seemingly easy tack of occulting /mu^{n}tu/ → **mu.^{n}tu* with /mu^{n}tu/ → *muː^{n}tu* is actually the wrong way to go, though, since the lengthening of the vowel cannot be explained. Rather, the right move is to map /mu^{n}tu/ unfaithfully onto *mu.tu*, simplifying the underlying prenasalized consonant. This reduces Luganda to a kind of chain-shift, with side-effects on vowel length: /nt/ → ^{n}*t* and /^{n}t/ → *t*. It is then a straightforward matter to apply the theories of chain-shifts reviewed in section 3.3.3. In fact, Gnanadesikan (1997: 139ff.) analyzes an exactly parallel case from Sanskrit, which she calls a coalescence "paradox".

31 These results were checked using the *otsoft* package created by Bruce Hayes. It is available for download via http://www.humnet.ucla.edu/humnet/linguistics/people/hayes.

32 This system does have some odd properties. Under some permutations, it will allow a contrast between moraic and non-moraic pre-consonantal consonants to emerge on the surface: /paTka/ → *paTka* vs. /patka/ → *pat.ka*. Lexical contrasts like this do not seem to occur, perhaps because the evidence for them is so very indirect. And under some permutations, this system will map (only) non-moraic pre-consonantal consonants onto zero: /paTka/ → *paTka* vs. /patka/ → *pa.ka*. This might offer a mora-based approach to certain kinds of ghost-segment behavior (cf. Clements and Keyser 1983, Zoll 1993).

33 I am greatly indebted to Elliott Moreton for discussion of this material.

REFERENCES

Abboud, Peter (1979). The verb in Northern Najdi Arabic. *Bulletin of the School of Oriental and African Studies* **42**. 467-499.

Abu-Mansour, Mahasen Hasan (1987). A Nonlinear Analysis of Arabic Syllabic Phonology, With Special Reference to Makkan. Doctoral dissertation, University of Florida, Gainesville, FL.

Alderete, John (1998). Morphologically-Governed Accent in Optimality Theory. Doctoral dissertation, University of Massachusetts, Amherst, Amherst, MA.

Alderete, John (1999). Faithfulness to prosodic heads. In Ben Hermans and Marc van Oostendorp (eds.) *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: John Benjamins.

Al-Mozainy, Hamza (1976). Vowel deletion and the segmental cycle in the Arabic dialect of Hijaz (Saudi Arabia). Master's thesis, University of Texas, Austin, Austin, TX.

Al-Mozainy, Hamza, Robert Bley-Vroman, and John McCarthy (1985). Stress shift and metrical structure. *Linguistic Inquiry* **16**. 135-144.

Al-Mozainy, Hamza Q. (1981). Vowel Alternations in a Bedouin Hijazi Arabic Dialect: Abstractness and Stress. Doctoral dissertation, University of Texas, Austin, Austin, TX.

Anderson, Stephen R. (1974). *The Organization of Phonology*. New York: Academic Press.

Aoun, Youssef (1979). Is the syllable or the supersyllable a constituent? *MIT Working Papers in Linguistics* **1**. 140-148.

Archangeli, Diana, and Douglas Pulleyblank (1994). *Grounded Phonology*. Cambridge, Massachusetts: MIT Press.

Bach, Emmon (1968). Two proposals concerning the simplicity metric in phonology. *Glossa*

2. 128-149.

Bakovic, Eric (1998). Elsewhere effects in Optimality Theory. Ms., Rutgers University, New Brunswick, NJ.

Bakovic, Eric (2000). Harmony, Dominance, and Control. Doctoral dissertation, Rutgers University, New Brunswick, NJ.

Beckman, Jill (1997). Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* **14**. 1-46.

Beckman, Jill (1998). Positional Faithfulness. Doctoral dissertation, University of Massachusetts, Amherst. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Benua, Laura (1997a). Affixes Classes are defined by Faithfulness. In Viola Miglio and Bruce Moren (eds.) *University of Maryland Working Papers in Linguistics 5. Selected Phonology Papers from Hopkins Optimality Theory Workshop 1997 / University of Maryland Mayfest 1997*. 1-27.

Benua, Laura (1997b). Transderivational Identity: Phonological Relations between Words. Doctoral dissertation, University of Massachusetts, Amherst. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Bermúdez-Otero, Ricardo (1999). Opacity: Evidence from West Germanic gemination. Handout of paper presented to Linguistic Association of Great Britain, Manchester.

Black, H. Andrew (1991). The phonology of the velar glide in Axininca Campa. *Phonology* **8**. 183-217.

Blanc, Haim (1970). The Arabic dialect of the Negev Bedouins. In *Proceedings of the Israel Academy of Sciences and Humanities*. Jerusalem: Israel Academy of Sciences and Humanities. 112-150.

Blevins, Juliette (1995). The syllable in phonological theory. In John A. Goldsmith (ed.) *The*

Handbook of Phonological Theory. Cambridge, MA, and Oxford, UK: Blackwell. 206-244.

Booij, Geert (1995). *The phonology of Dutch*. Oxford: Clarendon Press.

Booij, Geert (1996). Lexical phonology and the derivational residue. In Jacques Durand and Bernard Laks (eds.) *Current trends in phonology: Models and methods*. Salford, Manchester, UK: University of Salford. 69-96.

Booij, Geert (1997). Non-derivational phonology meets lexical phonology. In Iggy Roca (ed.) *Constraints and Derivations in Phonology*. Oxford: Oxford University Press. 261-288.

Borgstrøm, C. Hj. (1937). The Dialect of Barra in the Outer Hebrides. *Norsk Tidsskrift for Sprogvidenskap* **8**. 71-242.

Borgstrøm, C. Hj. (1940). *A Linguistic Survey of the Gaelic Dialects of Scotland*: Norsk Tidsskrift for Sprogvidenskap, Suppl. Bind 1.

Borowsky, Toni (1993). On the word level. In Sharon Hargus and Ellen Kaisse (eds.) *Studies in Lexical Phonology*. New York: Academic Press. 199-234.

Bosch, Anna (1991). Phonotactics at the Level of the Phonological Word. Doctoral dissertation, University of Chicago.

Bosch, Anna, and Ken de Jong (1998). Syllables and supersyllables: Evidence for low level phonological domains. Ms., University of Kentucky and Indiana University, Lexington, KY and Bloomington, IN.

Bosch, Anna, and Kenneth de Jong (1997). The prosody of Barra Gaelic epenthetic vowels. *Studies in the Linguistic Sciences* **27**. 1-15.

Broselow, Ellen (1992). Parametric variation in Arabic dialect phonology. In Ellen Broselow, M. Eid and John McCarthy (eds.) *Perspectives on Arabic Linguistics*. Amsterdam and Philadelphia: John Benjamins. 7-45.

Bye, Patrik (1996). Correspondence in the prosodic hierarchy and the grid: Case studies in overlength and level stress. Cand Philol Thesis, University of Tromso, Tromso, Norway.

Campbell, Lyle (1973). Extrinsic order lives, Indiana University Linguistics Club, Bloomington, IN.

Causley, Trisha (1997). Identity and featural correspondence: The Athapaskan case. In Kiyomi Kusumoto (ed.) *Proceedings of NELS 27*. Amherst, MA: GLSA. 93-105.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam, and Morris Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.

Clayton, Mary L. (1976). The redundance of underlying morpheme-structure conditions. *Language* **52**. 295-313.

Clements, G. N. (1986a). Compensatory lengthening and consonant gemination in Luganda. In L. Wetzels and E. Sezer (eds.) *Studies in Compensatory Lengthening*. Dordrecht: Foris. 37-77.

Clements, G. N. (1986b). Syllabification and epenthesis in the Barra dialect of Gaelic. In Koen Bogers, Harry van der Hulst and Maarten Mous (eds.) *The Phonological Representation of Suprasegmentals*. Dordrecht: Foris. 317-336.

Clements, G. N. (1997). Berber syllabification: derivations or constraints? In Iggy Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Oxford University Press. 289-330.

Clements, G. N., and Samuel Jay Keyser (1983). *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Cohn, Abigail, and John J. McCarthy (1994). Alignment and parallelism in Indonesian phonology, Cornell University and University of Massachusetts, Amherst. [Published

(1998) in *Working Papers of the Cornell Phonetics Laboratory* 12, 53–137.]

Crowhurst, Megan (1991). Demorafication in Tubatulabal: Evidence from initial reduplication and stress. In Tim Sherer (ed.) *Proceedings of NELS 21*. Amherst: GLSA. 49-63.

Davey, B. A., and H. A. Priestley (1990). *Introduction to Lattices and Order*. Cambridge: Cambridge University Press.

Dresher, B. Elan (1993). The chronology and status of Anglian smoothing. In Sharon Hargus and Ellen M. Kaisse (eds.) *Studies in Lexical Phonology*. San Diego, CA: Academic Press. 325-341.

Farwaneh, Samira (1995). Directionality Effects in Arabic Dialect Syllable Structure. Doctoral dissertation., University of Utah, Salt Lake City, UT.

Gafos, Adamantios (1996). The Articulatory Basis of Locality in Phonology. Doctoral dissertation, Johns Hopkins University, Baltimore, MD.

Gafos, Adamantios, and Linda Lombardi (1999). Consonant transparency and vowel echo. In Pius N. Tamanji, Mako Hirotani and Nancy Hall (eds.) *Proceedings of NELS 29, vol. 2: Papers from the Poster Sessions*. Amherst, MA: GLSA. 81-95.

Gafos, Diamandis (1998). Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* **16**. 223-278.

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

Gnanadesikan, Amalia (1997). Phonology with Ternary Scales. Doctoral dissertation, University of Massachusetts at Amherst, Amherst, MA.

Gordon, Matthew (1999). Syllable Weight: Phonetics, Phonology, and Typology. Doctoral dissertation, UCLA, Los Angeles, CA.

Green, Anthony Dubach (1997). The Prosodic Structure of Irish, Scots Gaelic, and Manx.

Doctoral dissertation, Cornell University, Ithaca, NY. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Guerssel, Mohammed (1986). Glides in Berber and Syllabicity. *Linguistic Inquiry* **17**. 1-12.

Hale, Mark, and Madelyn Kissock (1998). The phonology-syntax interface in Rotuman. In Matthew Pearson (ed.) *Recent Papers in Austronesian Linguistics: Proceedings of the Third and Fourth Meetings of the Austronesian Formal Linguistics Society, UCLA Occasional Papers in Linguistics #21*. Los Angeles: UCLA Department of Linguistics. 115-128.

Hale, Mark, Madelyn Kissock, and Charles Reiss (1998). Output-output correspondence in OT. In E. Curtis, J. Lyle and G. Webster (eds.) *Proceedings of WCCFL 16*. Stanford, CA: CSLI. 223-236.

Halle, Morris (1995). Letter commenting on Burzio's article. *Glot International* **1(9/10)**.

Halle, Morris, and William Idsardi (1997). *r*, hypercorrection, and the Elsewhere Condition. In Iggy Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Clarendon Press. 331-348.

Harris, James (1987). Sonority and syllabification in Spanish. Ms., MIT, Cambridge, MA.

Harris, James (1993). Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan. In Kenneth Hale and Samuel Jay Keyser (eds.) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. 177-194.

Harris, John (1990). Derived phonological contrasts. In Susan Ramsaran (ed.) *Studies in the Pronunciation of English*. London: Routledge. 87-105.

Hastings, Ashley (1974). *Stifling*. Bloomington, IN: Indiana University Linguistics Club.

Hayes, Bruce (1986). Inalterability in CV phonology. *Language* **62**. 321-251.

Hayes, Bruce (1987). A revised parametric metrical theory. In J. McDonough and B. Plunkett (eds.) *Proceedings of NELS 17*. Amherst: GLSA. 274-289.

Hayes, Bruce (1989). Compensatory Lengthening in moraic phonology. *Linguistic Inquiry* **20**. 253-306.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Hogg, Richard M. (1978). The Duke of York gambit: A variation. *Lingua* **44**. 255-266.

Hung, Henrietta (1994). The Rhythmic and Prosodic Organization of Edge Constituents. Doctoral dissertation, Brandeis University, Waltham, MA.

Hyman, Larry (1985). *A Theory of Phonological Weight*. Dordrecht: Foris.

Idsardi, William (1998). Tiberian Hebrew spirantization and phonological derivations. *Linguistic Inquiry* **29**. 37-73.

Ingham, Bruce (1982). *Northeast Arabian Dialects*. London: Kegan Paul International.

Inkelas, Sharon (1999). Exceptional stress-attracting suffixes in Turkish: representation vs. the grammar. In René Kager, Harry van der Hulst and W. Zonneveld (eds.) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press. 134-187.

Irshied, Omar (1984). The Phonology of Bani-Hassan Arabic, a Bedouin Jordanian dialect. Doctoral dissertation, University of Illinois at Champaign-Urbana, Champaign, IL.

Irshied, Omar, and Michael Kenstowicz (1984). Some phonological rules of Bani-Hassan Arabic, a Bedouin dialect. *Studies in the Linguistic Sciences* **14**. 109-147.

Ito, Junko, Yoshihisa Kitagawa, and Armin Mester (1996). Prosodic faithfulness and correspondence: Evidence from a Japanese Argot. *Journal of East Asian Linguistics* **5**. 217-294.

Iverson, Gregory K. (1995). Rule ordering. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge, MA, and Oxford, UK: Blackwell. 609-614.

Johnstone, T. M. (1967a). Aspects of syllabication in the spoken Arabic of `Anaiza. *Bulletin* *of the School of Oriental and African Studies* **30(1)**. 1-16.

Johnstone, T. M. (1967b). *Eastern Arabian dialect studies*. Oxford: Oxford University Press.

Kager, Rene (1997). Rhythmic vowel deletion in Optimality Theory. In Iggy Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Oxford University Press. 463-499.

Keer, Edward (1999). Geminates, the OCP, and Faithfulness. Doctoral dissertation, Rutgers University, New Brunswick, NJ. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, Michael (1995). Cyclic vs. non-cyclic constraint evaluation. *Phonology* **12**. 397-436.

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

Kenstowicz, Michael, and Charles Kisseberth (1979). *Generative Phonology: Description and Theory*. New York: Academic Press.

Kiparsky, Paul (1973a). 'Elsewhere' in phonology. In Stephen R. Anderson and Paul Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston. 93-106.

Kiparsky, Paul (1973b). Phonological representations. In O. Fujimura (ed.) *Three Dimensions of Linguistic Theory*. Tokyo: TEC. 3-136.

Kiparsky, Paul (1997a). Covert generalization. In Geert Booij, Angela Ralli and Sergio Scalise (eds.) *Proceedings of the First Mediterranean Conference of Morphology (Mytilene, Greece, Sept. 19-21 1997)*: University of Patras. 65-76.

Kiparsky, Paul (1997b). LP and OT. Handout of paper presented to LSA Summer Linguistic Institute, Cornell University.

Kiparsky, Paul (2000). *Paradigmatic Effects*. Stanford, CA: CSLI.

Kirchner, Robert (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* **27**. 341-350.

Klokeid, Terry (1977). Some irrelevant observations concerning rule interaction. *Journal of Linguistics* **13**. 283-285.

Koutsoudas, Andreas, Gerald Sanders, and Craig Noll (1974). On the application of phonological rules. *Language* **50**. 1-28.

de Lacy, Paul (1997). Prosodic Categorisation. Master's thesis, University of Auckland, Auckland, New Zealand.

Lamontagne, Greg, and Keren Rice (1995). A correspondence account of coalescence. In Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey (eds.) *University of Massachusetts Occasional Papers in Linguistics 18*. Amherst, MA: GLSA. 211-224.

Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Levin, Juliette (1985). A Metrical Theory of Syllabicity. Doctoral dissertation, MIT, Cambridge, Massachusetts.

Lombardi, Linda (1991). Laryngeal Features and Laryngeal Neutralization. Doctoral dissertation, University of Massachusetts, Amherst.

Łubowicz, Anna (1999). Derived environment effects in OT. Ms., University of Massachusetts, Amherst, Amherst, MA.

Marantz, Alec (1982). Re Reduplication. *Linguistic Inquiry* **13**. 483-545.

Mascaró, Joan (1984). Continuant spreading in Basque, Catalan, and Spanish. In Mark Aronoff and Richard Oehrle (eds.) *Language Sound Structure: Studies Presented to Morris Halle by his Teacher and Students*. Cambridge, MA: MIT Press. 287-298.

McCarthy, John (1986). OCP Effects: Gemination and antigemination. *Linguistic Inquiry* **17**. 207-263.

McCarthy, John (1991). Synchronic Rule Inversion. In L. Sutton, C. Johnson and R. Shields (eds.) *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Soceity*. Berkeley, CA: Berkeley Linguistics Society. 192-207.

McCarthy, John (1993). A case of surface constraint violation. *Canadian Journal of Linguistics* **38**. 169-195.

McCarthy, John (1995). Extensions of faithfulness: Rotuman revisited, University of Massachusetts, Amherst. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

McCarthy, John (1998). Sympathy and phonological opacity. Ms., University of Massachusetts, Amherst, Amherst, MA. [Presented at Hopkins Optimality Theory Workshop/University of Maryland Mayfest 1997. Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

McCarthy, John (1999a). Morpheme structure constraints and paradigm occultation. In M. Catherine Gruber, Derrick Higgins, Kenneth Olson and Tamra Wysocki (eds.) *CLS 32, vol. II: The panels*. Chicago, IL: Chicago Linguistic Society.

McCarthy, John (1999b). Sympathy and phonological opacity. *Phonology* **16(3)**. xxx-xxx.

McCarthy, John (2000). Harmonic serialism and parallelism. In Masako Hirotani (ed.) *Proceedings of the North East Linguistics Society 30*. Amherst, MA: GLSA. xxx-xxx. [Can be downloaded from http://www-unix.oit.umass.edu/~jjmccart/.]

McCarthy, John (to appear). The prosody of phase in Rotuman. *Natural Language and Linguistic Theory*. [Can be downloaded from http://www-unix.oit.umass.edu/~jjmccart/.]

McCarthy, John, and Alan Prince (1986). Prosodic Morphology, University of Massachusetts at Amherst and Brandeis University.

McCarthy, John, and Alan Prince (1988). Quantitative transfer in reduplicative and templatic morphology. In Linguistic Society of Korea (ed.) *Linguistics in the Morning Calm*.

Seoul: Hanshin Publishing Co. 3-35.

McCarthy, John, and Alan Prince (1990a). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* **8**. 209-283.

McCarthy, John, and Alan Prince (1990b). Prosodic morphology and templatic morphology. In Mushira Eid and John McCarthy (eds.) *Perspectives on Arabic linguistics II: Papers from the Second Annual Symposium on Arabic Linguistics*. Amsterdam: John Benjamins. 1-54.

McCarthy, John, and Alan Prince (1993a). Generalized Alignment. In Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology*. Dordrecht: Kluwer. 79-153. [Excerpts reprinted in in John Goldsmith, ed., *Essential Readings in Phonology*. Oxford: Blackwell. Pp. 102–136, 1999.]

McCarthy, John, and Alan Prince (1993b). Prosodic Morphology I: Constraint Interaction and Satisfaction, University of Massachusetts, Amherst, and Rutgers University. Report RUCCS TR-3. [To appear, MIT Press, Cambridge, MA.]

McCarthy, John, and Alan Prince (1995). Faithfulness and Reduplicative Identity. In Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey (eds.) *University of Massachusetts Occasional Papers in Linguistics UMOP 18*. Amherst, MA: GLSA. 249-384.

McCarthy, John, and Alan Prince (1999). Faithfulness and identity in Prosodic Morphology. In Rene Kager, Harry van der Hulst and Wim Zonneveld (eds.) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press. 218-309.

Merchant, Jason (1997). Sympathetic devoicing and continuancy in Catalan. *Phonology at Santa Cruz* **5**. 57-62.

Mitchell, T. F. (1960). Prominence and syllabication in Arabic. *Bulletin of the School of Oriental and African Studies* **23**. 369-389. [Reprinted in Mitchell, T. F. 1975

Principles of Firthian Linguistics. London: Longmans. Also reprinted in Al-Ani, Salman. 1978 *Readings in Arabic Linguistics*. Bloomington, IN: Indiana University Linguistics Club.]

Mohanan, K. P. (1985). Syllable structure and lexical strata in English. *Phonology* **2**. 139-155.

Moravcsik, Edith (1978). Reduplicative constructions. In Joseph Greenberg (ed.) *Universals of Human Language*. Stanford: Stanford University Press. 297-334.

Moreton, Elliott (1996/1999). Non-computable functions in Optimality Theory. Ms., University of Massachusetts, Amherst, Amherst, MA. [Revised and available (1999) on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html. Originally circulated in 1996.]

Myers, Scott (1987). Vowel shortening in English. *Natural Language and Linguistic Theory* **5**. 485-518.

Ní Chiosáin, Máire, and Jaye Padgett (1997). Markedness, segment realization, and locality in spreading, UC Santa Cruz, Santa Cruz, CA. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Noyer, Rolf (1997). Attic Greek accentuation and intermediate derivational representations. In Iggy Roca (ed.) *Constraints and Derivations in Phonology*. Oxford: Oxford University Press. 501-528.

Paradis, Carole (1997). Non-transparent constraint effects in Gere: from cycles to derivations. In Iggy Roca (ed.) *Constraints and Derivations in Phonology*. Oxford: Oxford University Press. 529-550.

Pater, Joe (1995). On the nonuniformity of weight-to-stress and stress preservation effects in English, McGill University. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Pater, Joe (1999). Austronesian nasal substitution and other NC effects. In Rene Kager, Harry van der Hulst and Wim Zonneveld (eds.) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press. 310-343.

Piggott, G. L. (1995). Epenthesis and syllable weight. *Natural Language and Linguistic Theory* **13**. 283-326.

Potter, Brian (1994). Serial optimality in Mohawk prosody. In Katharine Beals, Jeannette Denton, Robert Knippen, Lynette Melmar, Hisami Suzuki and Erica Zeinfeld (eds.) *Proceedings of the Thirtieth Annual Regional Meeting of the Chicago Linguistics Society*. Chicago, IL: Chicago Linguistics Society. 347-361.

Prince, Alan (1975). The Phonology and Morphology of Tiberian Hebrew. Doctoral dissertation, MIT, Cambridge, MA.

Prince, Alan (1990). Quantitative consequences of rhythmic organization. In M. Ziolkowski, M. Noske and K. Deaton (eds.) *Parasession on the Syllable in Phonetics and Phonology*. Chicago: Chicago Linguistic Society. 355-398.

Prince, Alan (1996). A letter from Alan Prince. *Glot International* **6(2)**.

Prince, Alan (1997). Elsewhere & otherwise, Rutgers University. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Prince, Alan, and Paul Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar, Rutgers University, New Brunswick, NJ. Report RUCCS TR-2. [To appear, MIT Press, Cambridge, MA.]

Prochazka, Theodore (1988). *Saudi Arabian dialects*. London: Kegan Paul International.

Pulleyblank, Douglas (1996). Neutral vowels in Optimality Theory: A comparison of Yoruba and Wolof. *Canadian Journal of Linguistics* **41**. 295-347.

Pullum, Geoffrey (1976). The Duke of York gambit. *Journal of Linguistics* **12**. 83-102.

Roca, Iggy (1997). Derivations or constraints, or derivations and constraints? In Iggy Roca (ed.) *Constraints and Derivations in Phonology*. Oxford: Oxford University Press. 3-42.

Rosenthall, Samuel (1994). Vowel/Glide Alternation in a Theory of Constraint Interaction. Doctoral dissertation, University of Massachusetts, Amherst, Amherst, MA.

Rubach, Jerzy (1984). *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris.

Rubach, Jerzy (1993). *The Lexical Phonology of Slovak*. Oxford: Oxford University Press.

Rubach, Jerzy (1997). Extrasyllabic consonants in Polish: derivational Optimality Theory. In Iggy Roca (ed.) *Constraints and Derivations in Phonology*. Oxford: Oxford University Press. 551-582.

Rubach, Jerzy (2000). Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* **31**. 271-317.

Samek-Lodovici, Vieri (1992). Universal constraints and morphological gemination: A crosslinguistic study. Ms., Brandeis University, Waltham, MA.

Samek-Lodovici, Vieri, and Alan Prince (1999). Optima. Ms., University of London and Rutgers University, New Brunswick, NJ. [Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.]

Sanders, Gerald (1974). Precedence relations in language. *Foundations of Language* **11**. 361-400.

Sapir, Edward, and Morris Swadesh (1978). *Nootka texts: Tales and ethnological narratives, with grammatical notes and lexical material*. New York: AMS Press.

Scobbie, James (1993). Constraint violation and conflict from the perspective of Declarative Phonology. *Canadian Journal of Linguistcs/Revue canadienne de linguistique* **38**. 155-167. [Special issue edited by Darlene LaCharité and Carole Paradis.]

Selkirk, Elisabeth O. (1981). Epenthesis and degenerate syllables in Cairene Arabic. In Hagit

Borer and Joseph Aoun (eds.) *Theoretical Issues in the Grammar of the Semitic Languages*. Cambridge, MA: MIT. 111-140.

Sherer, Tim (1994). Prosodic Phonotactics. Doctoral dissertation, University of Massachusetts, Amherst, Amherst, MA.

Smolensky, Paul (1995). On the internal structure of the constraint component Con of UG. Handout of paper presented to talk, University of Arizona.

Walker, Rachel (1998). Nasalization, Neutral Segments, and Opacity Effects. Doctoral dissertation, University of California, Santa Cruz, Santa Cruz, CA.

Wilson, Colin (1999). Consonant cluster neutralization and targeted constraints. Ms., Johns Hopkins University, Baltimore, MD.

Wiltshire, Caroline (1992). Syllabification and Rule Application in Harmonic Phonology. Doctoral dissertation, University of Chicago.

Zec, Draga (1995). Sonority constraints on syllable structure. *Phonology* **12**. 85-129.

Zoll, Cheryl (1993). Ghost segments and Optimality. In Erin Duncan, Michele Hart and Philip Spaelti (eds.) *Proceedings of WCCFL 12*. Stanford, CA: Stanford Linguistics Association. 183-199.

Zwicky, Arnold M. (1974). Taking a false step. *Language* **50**. 215-224.