SYMPATHY, CUMULATIVITY, AND THE DUKE-OF-YORK GAMBIT[*]

John J. McCarthy
University of Massachusetts, Amherst

Abstract: The Duke-of-York gambit (Pullum 1976) involves derivations of the form A→B→A, where underlying /A/ passes through an intermediate stage B before returning to A at the surface. Optimality Theory (Prince and Smolensky 1993) has significant implications for the Duke of York gambit. Furthermore, attested and unattested Duke-of-York cases have implications for the analysis of opacity in Optimality Theory using sympathy (McCarthy 1998; to appear). A key idea is that derivations must be cumulative, and a measure of cumulativity is incorporated into sympathy theory.

1. Introduction

Serial derivations have been a central idea in generative phonology since the beginning, but scant attention has been paid to a key question: is any serial derivation possible in human languages? More precisely, can any 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. Research in generative phonology has largely focused instead on the form of rules and the nature of representations — subjects which are interesting in themselves but do not much 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, some analyses of *r*-dropping and intrusion in English dialects work this way, first deleting final *r* and then re-inserting it before a vowel: *Homer is* → *Hom[ə] is* → *Homer is*. 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. He concludes that linguistic theory does need to provide DY power.

Optimality Theory (Prince and Smolensky 1993) offers a very different 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. In this article, I 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/	/CBD/
⇨	A→B/__D	CBD		ZBD	—
⇨	B→A/C__	CAD		—	CAD

The DY interaction is vacuous here 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, as I will show below.

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.[1]

DY derivations of the vacuous type are abundantly attested; indeed, all of Pullum's examples are like this, as are 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; indeed, Pullum cites no cases at all. Two possible examples, from Tiberian Hebrew and Bedouin Arabic, are discussed and reanalyzed in section 3. The conclusion I reach is that, except for "cyclic" effects (which can be modeled with output-output faithfulness), feeding DY derivations do not exist. This typological result demands an explanation, and in the following sections of this article I seek to supply one.

One element of the explanation is sympathy theory (McCarthy 1998; to appear), discussed below 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. Sympathy appears to be sufficient to account

---

[1] A third type of DY derivation involves a *bleeding* interaction at the intermediate stage — A→B in order to wait out some process, and then B→A:

(i) Bleeding DY Derivation

	Underlying	/CAD/	
	A→B/__D	CBD	A→B to escape next rule
	C→E/__A	*dna*	No A there to condition C→E change.
	B→A/__D	CAD	B→A, undoing effect of first rule.

A case of this type occurs in Catalan (Harris 1993; Merchant 1997).

for all attested interactions that are subsumed under Kiparsky's (1973b) definition of opacity.

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 notion of 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 bears further study, and a first installment of that is provided in section 6. DY derivations are only the most obvious situations where cumulativity is breached. Any derivation where an underspecified property, such as syllabification, is filled-in and later altered also looks non-cumulative. 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.

Clearly, then, processes of syllabification and resyllabification are relevant to understanding the cumulativity property. Below, I conclude 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 are fully compatible with the revised theory of sympathy.

2. Vacuous Duke-of-York Derivations

There is no shortage of real DY derivations like (1). Some examples, most of which were originally collected by Pullum (1976), appear in (4):

(4) Vacuous DY Cases
     a. Nootka rounding/unrounding (Campbell 1973).
     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; 1993a).
     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).

These examples all share certain characteristic properties: there are two (or more) rules which produce opposite mappings (A→B and B→A); these rules apply in environments which sometimes overlap; and the rules are ordered with the A→B rule applying before the B→A rule. This constellation of properties will produce 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, exemplified in (5c), where a dorsal consonant is both preceded by a round vowel and followed by a syllable boundary, so it meets the

structural conditions of both rules. In this case, Delabialization takes precedence:

(5) Nootka (De)labialization (Campbell 1973; Pullum 1976; Sapir and Swadesh 1978)
    a. Dorsals become labialized after round vowels

        K → K^w / o___   ʔo**k**^{w}iː&#x141;   'making it'
                                  cf. **k**iː&#x141;   'making'

    b. Syllable-final labiodorsals delabialize

        K^w → K / ___.   &#x141;aːk̲.ši(λ)   'to take pity on'
                                  cf. &#x141;aː.**k**^{w}iqnak 'pitiful'

    c. Interaction: Delabialization wins

                                ṁoː**q**.   'throwing off sparks'
                                cf. ṁo.**q**^{w}ak   'phosporescent'

Under strict serialism, the only way to ensure that Delabialization wins is to let it have the last word, by ordering it after Labialization. The result is a DY derivation in just those cases where the ordering matters, such as input /ṁoːq/:

(6) Serial Derivation for Nootka

	Underlying	/ṁoːq/	cf.	/ʔokiː&#x141;/	/&#x141;aːk^{w}ši(λ)/
⇨	Labialization	ṁoːq^w.		ʔo.k^{w}iː&#x141;	—
⇨	Delabialization	ṁoːq.		—	&#x141;aːk.ši(λ)

Nootka, then, has exactly the characteristics of a vacuous DY derivation: two rules which produce contradictory mappings in environments that sometimes overlap.

    Before we continue, it is necessary to consider and dismiss two alternatives that might seem like reasonable ways to side-step the whole 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, Sanders, and Noll 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:

(7) Nootka with Disjunctive Ordering

Underlying	/ṁoːq/	cf.	/ʔokiː&#x141;/	/&#x141;aːk^{w}ši(λ)/
Delabialization	ṁoːq.[2]		—	&#x141;aːk.ši(λ)
Labialization	*blocked by EC*		ʔo.k^{w}iː&#x141;	—

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 in principle and in fact eliminate only a

---

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

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.[3]

Another failed approach to Nootka would involve skirting the intermediate stage of the /ṁoːq/ → ṁoːqʷ → ṁoːq derivation by enforcing the effect of Labialization in the underlying representation: /ṁoːqʷ/ → ṁoːq. To handle tautomorphemic *o*K sequences like /ṁoːqʷ/, Labialization functions as a morpheme structure constraint. But heteromorphemically, as in (5a), Labialization functions as a regular rule. This is just the familiar 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" Prince and Smolensky 1993).

With these failed alternatives aside, we turn to OT. 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 (8a). The other prohibits rounded dorsals syllable-finally (8b):

(8) Constraints for Nootka
     a. "Rounding"
           *oK
     b. "Unrounding"
           *K^w]$_\sigma$

To be visibly active, these markedness constraints must dominate an appropriate faithfulness constraint, IDENT(round). These rankings are responsible for the alternations in (5a, b):[4]

(9) "Rounding" >> IDENT(round)

/ʔokiːɬ/	"Rounding"	IDENT(round)
☞ (a) ʔo.kʷiːɬ		*
(b) ʔo.kiːɬ	* !	

---

[3]The Bedouin Arabic vowel-height alternations are similarly problematic: 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. Clearly, there is no specific/general relation between these rules either.

[4]In a full analysis, it would also be necessary to consider failed 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.

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

/ƛaːkʷši(ƛ)/	"UNROUNDING"	IDENT(round)
☞ (a) ƛaːk.ši(ƛ)		*
(b) ƛaːkʷ.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", the output is unrounded in situations of conflict like /m̓oːq/.

(11) "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. Moreover, with faithfulness bottom-ranked, the choice of input — /m̓oːq/, /m̓oːqʷ/, or archisegmental /m̓oːQ/ — doesn't matter, since all map to surface *m̓oːq*. So there is no need to restrict the inputs and no Duplication Problem. Finally, as usual in OT, ranking permutation predicts a range of permitted interlinguistic variation. In particular, if the ranking of the two markedness constraints were reversed, then *m̓oːqʷ* would be the output.

All of the cases of vacuous DY derivations 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 rules must be explained in terms of 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.[5] By decoupling precedence from serial ordering, OT permits the vacuous DY cases in (4) to be analyzed without positing the spurious intermediate stage.

We have seen, then, 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.

---

[5]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).

3. Feeding Duke-of-York Derivations

In feeding DY derivations, the intermediate stage is crucial. The A→B rule feeds some process which applies at the intermediate stage, before the B→A rule wipes out its environment. Seemingly plausible hypothetical examples are not difficult to construct. The first, given in (12), is modeled after a post-vocalic spirantization process in Tiberian Hebrew, but with a twist: a general process inserting *ə* after codas applies before spirantization, and then the *ə*'s are deleted after spirantization:

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

Another hypothetical example was invented by Paul Kiparsky (e-mail communication, 7 July 1998). At the first step, trimoraic CVVC syllables are repaired by *i* epenthesis. 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 CVVC syllable is re-repaired by shortening. Because it shares some rules with the real Yokuts language, I will call this hypothetical system quasi-Yokuts:

(13) Quasi-Yokuts
	Underlying	/maat/
⇨	Epenthesis	maati	To repair trimoraic syllable.
	Palatalization	maači	*ti → či* generally.
⇨	Apocope	maač	Final vowels delete
	Shortening	mač	To repair trimoraic syllable.

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 every rule is a completely natural process for which precedents can easily be found in real languages. The peculiar thing is not the rules themselves, but their co-existence and interaction in a single system.

And indeed these cases are peculiar. Examples don't exactly lie thick on the ground, or even thin. Pullum's (1976) survey contains none, and I am aware of just two putative cases, one in real Tiberian Hebrew (Prince 1975) and the other in Bedouin Arabic (Al-Mozainy 1981).[6] The Hebrew example will be dispensed with fairly quickly — it turns out to be a cyclic or "output-output" faithfulness effect. I will then 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 (output-output effects excepted) do not in general occur — an observation for which linguistic theory needs to supply an explanation.

---

[6]Some approaches to neutral vowels in vowel harmony systems supply a third potential DY case. Bach (1968) first proposed that neutral vowels actually undergo the harmony process at an intermediate stage of the derivation, only to be shifted back in surface representation: e.g., Finnish /tuoli-llA/ → *tuoli̇lla* → *tuolilla* 'on the chair'. (See Ní Chiosáin and Padgett (1997) and Walker (1998) for recent discussion from an OT perspective.) Of course, autosegmental phonology supplies alternative, structural interpretations of neutral segments that require no DY derivation. The issue is too big to address here, but see van der Hulst and van der Weijer (1995) for a recent review.

First the (real) Hebrew case. Non-finite verbs prefixed with *bi-* 'in' and *li-* 'infinitive' systematically differ in whether or not the second consonant of the root undergoes post-vocalic spirantization: [bixθōβ] 'when writing' vs. [lixtōβ] 'to write'. Prince (1975) proposes that the stem is /ktob/, and that spirantization follows indirectly from a difference in junctural strength between these prefixes, which leads to a (temporary) difference in ə epenthesis:[7]

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

	Underlying	/ba#ktob/	/la+ktob/
⇨	ə Epenthesis	ba#kətob	—
	Spirantization	ba#xəθoβ	laxtoβ
⇨	ə Deletion	ba#xθoβ	—
	Other rules	bixθōβ	lixtōβ
		'when writing'	'to write'

What distinguishes this case from quasi-Hebrew in (12) is the crucial role played by the morphology. In quasi-Hebrew, the feeding DY derivation is indifferent to the morphology, but in the real case, the morphology is decisive. As Prince indicates with the # and + boundaries, there is a difference in behavior between stems prefixed with the preposition *bi –* and stems prefixed with the inflectional morpheme *li –*; only the former have ə Epenthesis, which is the prelude to the DY derivation.

In OT, the spirantized θ in *bixθōβ* can be analyzed as an effect of faithfulness to the free-standing word *kəθōβ* 'writing' through the output-output constraint IDENT(cont). The difference in junctural strength between *bi-* and *li-* supports this approach: paralleling the English Level II/Level I distinction, it makes sense to say that these affixes induce distinct 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.

Now on to Bedouin Arabic. Analyses of this language posit a feeding DY derivation that cannot be reanalyzed in terms of output-output faithfulness. To understand the DY derivation and my reanalysis of it, it's first necessary to have a good deal of background in Bedouin Arabic phonology. My plan is to first present the core processes of vowel raising and deletion, and then turn to their interaction with stress, which is where DY comes in.

The following partial paradigms illustrate the main points:[8]

---

[7]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.

[8]The transcription has been simplified by suppressing indications of velar palatalization and contextual effects on the raised vowel. The data are drawn from the Ḥarbi dialect carefully described and analyzed by Al-Mozainy (1976; Al-Mozainy, Bley-Vroman, and McCarthy 1985; 1981)}. (There is in addition a text dictated by Al-Mozainy in Ingham (1982, 112-115).) Quite 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 Cairene or Palestinian, are very different indeed, reflecting an ancient division between two dialect groups.

(15) 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/).

Al-Mozainy (1981) takes considerable care in establishing that the underlying forms given in (15) are correct. The most controversial feature is the vowel posited in the first syllable of /katab/ — allegedly underlying /a/, but always *i* or Ø on the surface. Three distinct 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: after a guttural consonant or before *a* if the intervening consonant is a guttural or a coronal sonorant (Al-Mozainy 1981; Gafos and Lombardi In preparation; Irshied and Kenstowicz 1984; McCarthy 1994). Some examples of the blocking effect, showing preservation of underlying /a/: /saħab/ → *saħab* 'he pulled'; /daras/ → *daras* 'he studied'; /ʔakal/ → *ʔakal* 'he ate'. 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. Examples discussed by Al-Mozainy include a natural consonant-permuting secret language (/dafaʕ/ → *difaʕ/daʕaf/ʕadaf/faʕad/fidaʕ...* 'he pushed'; /ðarab/ → *ðarab/ðibar/ribað/barað...* 'he beat') and an artificial *ħar*-inserting one (/katab/ → *kitab/kaħartab* 'he wrote').

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 (15):

(16) Core Rules for Bedouin Arabic
   a. Syncope
       i → Ø / ___]$_\sigma$ σ   "Delete short *i* in a (non-final) light syllable."
   b. Trisyllabic Deletion
       V → Ø /___]$_\sigma$ L σ   "Delete a short vowel from an open syllable that is followed by a (non-final) light syllable."
   c. Raising
       a → i / ___]$_\sigma$ σ   "Raise *a* to *i* in a (non-final) open syllable."
   d. Stress
       Latin stress rule: final syllable extrametrical, moraic trochees R→L.

The rule of Syncope is necessary to account for alternations like *samiʕ/samʕat*. Raising is exemplified by forms like *kitab*; as was just noted, Raising is blocked under various circumstances. The stress appears to follow the Latin model, as in many of the sedentary Arabic dialects: stress the penult if heavy,

otherwise the antepenult.[9]

The real complication in this rule system is obviously Trisyllabic Deletion. As stated, it can affect any short vowel, though in practice (because it is ordered after Syncope), it only affects /a/. The targeted vowel must be in a light syllable which is itself followed by a non-final light syllable — a baroque non-local condition that suggests something is being missed here. Stay tuned.

These rules have several crucial ordering relations, illustrated in the following derivations:

(17) Rule Interaction

	/katab/	/katab-at/	/samiʕ/	/samiʕ-at/	/kitib/	/kitib-at/
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 $i \rightarrow \varnothing / \_\_]_\sigma \sigma$				sam.ʕat	ktib	kit.bat
Trisyllabic Deletion $V \rightarrow \varnothing / \_\_\_]_\sigma L \sigma$		kta.bat				
Raising $a \rightarrow i / \_\_\_]_\sigma \sigma$	ki.tab	kti.bat	si.miʕ			
Stress (Latin rule)	kí.tab	ktí.bat	sí.miʕ	sám.ʕat	ktíb	kít.bat

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.) The data are consistent with late application of the Stress rule, though other evidence, to be discussed below, appears to require early stressing with subsequent stress shift (Al-Mozainy, Bley-Vroman, and McCarthy 1985).

When we come to consider the same facts in Optimality Theory, the first matter we must address is the /a/ → *i*, /i/ → Ø chain shift (Kirchner 1996; McCarthy 1993b). The insight behind the analysis of chain shifts in OT is relative faithfulness (Gnanadesikan 1997; Kirchner 1996; Prince 1998): 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.[10] 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

---

[9]As usual in Arabic, final superheavy syllables act like heavy penults, and so receive stress: *maktúːb* 'written', *saħábt* 'I pulled'.

[10]"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.

phonological scales, such as consonantal stricture, voicing and sonorancy, or vowel height.[11] All scales are ternary, by her hypothesis, and positions on a scale can be referenced by markedness and faithfulness constraints. The family of IDENT constraints (McCarthy and Prince 1995) is expanded to distinguish categorically between degrees of difference on a scale:

(18) IDENT and IDENT-ADJ
    a. IDENT(S)
           Corresponding segments have identical values on the scale S.
    b. IDENT-ADJ(S)
           Corresponding segments have identical *or adjacent* values 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 IDENT(S), just like the familiar IDENT for binary features. But the mapping /A/ → C also incurs a violation of IDENT-ADJ(S), and so it is categorically less faithful than the "closer" mappings.

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

(19) The Duration Scale *Dur*
    a   >   *i*   >   Ø

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

(20) REDUCE (Kirchner 1996, 347)
    Minimize the duration of a short vowel in an open syllable

The vowel *i* receives one mark from REDUCE, while *a* gets two. Ranked between IDENT-ADJ(Dur) and IDENT(Dur), REDUCE is responsible for the /a/ → *i* and /i/ → Ø chain shift, and it is correctly unable to compel the /a/ → Ø mapping:[12]

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

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

[11]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.

[12]I am ignoring a technical problem: since IDENT refers to segmental correspondents, it is irrelevant in /i/ → Ø mappings. Various solutions are possible, involving either featural correspondence or sub-moraic weight units (Bye 1999).

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

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

In (21), 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, (21c), but the cost is too high: fatal violation of top-ranked IDENT-ADJ(Dur), which bars the /a/ → Ø mapping. In (22), on the other hand, perfect performance on REDUCE is possible: the /i/ → Ø mapping only violates the low-ranking faithfulness constraint IDENT(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.

We now have sufficient background to permit study of Trisyllabic Deletion and its role in a DY derivation. The problem centers around the interaction of Trisyllabic Deletion with Stress. Though transparent Latin-type stress is the norm (23a), the effects of Trisyllabic Deletion produce some anomalies (23b):

(23) Stress Data[13]
    a. Words Consistent with a Latin-Type Trochaic Stress Rule
        i. Monosyllables
            ktíb
        ii. Disyllables
            kítab
            ktíbat
            sámʕat
        iii. Polysyllables with Heavy Penults
            maktúːfah   'tied (f. sg.)'
        iv. Polysyllables with Light Penults and Heavy Antepenults
            máːlana   'our property'
            ḍarábtukum   'I hit you (m. pl.)'
            yšuːfúːnukum   'they (m.) see you (f. pl.)'
    b. Words Inconsistent with a Latin-Type Trochaic Stress Rule
        i. Polysyllables with Light Penults and Antepenults (always with initial ʔ)
            ʔakálat   'she ate'   /ʔakal–at/
            ʔaxádat   'she took'   /ʔaxad–at/
        ii. Polysyllables with Trisyllabic Deletion
            ʔinksárat   'it (f.) got broken'   /ʔinkasar–at/
            ʕallmítuh   'she taught him'   /ʕallam–at-uh/

The words in (23a) have a stress pattern that could be derived by late application of a Latin-type trochaic stress rule, as in (17). But the words in (23b) do not, and they have been taken as evidence for early

---

[13]On final superheavy syllables, see footnote 9.

stressing with subsequent stress shift under deletion (Al-Mozainy, Bley-Vroman, and McCarthy 1985; Al-Mozainy 1981; Hayes 1995, 228–38). The idea is that Stress applies prior to Trisyllabic Deletion, which then sometimes deletes the stressed vowel. Deletion of a vowel out from under the stress forces stress to shift to the other syllable of the foot:

(24) Serial Analysis of Stress Anomalies

	a. /katab-at/	b. /ʔinkasar-at/	c. /ʔakal-at/
Trochaic 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

The derivation in (24a) is provided for comparison purposes; the interesting cases are (24b) and (24c). In (24b), 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 (24c), 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 specimen of the feeding DY species, because the intermediate stage is crucial to obtaining the stress-shift effect.

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? This situation is strongly reminiscent of trisyllabic shortening in English. Pairs like *serene/serenity*, *grateful/gratitude*, and *derive/derivative* 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.

The complexity and non-locality of English trisyllabic shortening have been addressed by incorporating reference to 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: $se[réni]_{Ft} <ty>$. In a sense made precise by Prince (1990), shortening improves the well-formedness of the trochaic foot, replacing a HL (heavy-light) trochee with a more harmonic LL trochee. This approach has answers to the *why*'s of trisyllabic shortening, as Prince (1996) has recently emphasized. 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, the key to understanding trisyllabic deletion in Bedouin Arabic is placing it in the context of an *iambic* stress system. 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 this reason, many languages have iambic lengthening processes, which improve LL iambs by lengthening the second syllable. Another logically possible consequence of the I/TL or GH 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.

The Arabian Bedouin Arabic dialects have been traditionally 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 straightforwardly iambic. In the course of reanalyzing the trisyllabic deletion phenomenon, I will show that Al-Mozainy's Arabian dialect is also iambic, though perhaps not obviously so.

The DY 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 seemingly trochaic data in (23a) are also compatible with iambic stress. To start with, the forms with a heavy penult in (23a.iii) can easily be given an iambic analysis instead of a trochaic one: *(kitáb)na*. With disyllables like (23a.ii), expected iambic stress — e.g., **(kitáb)* — gives way to NON-FINALITY, which is a near-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). Monosyllables like (23a.i) are, of course, equally compatible with iambic or trochaic stress, leaving only words with a light penult and a stressed heavy antepenult like (23a.iv) to be accounted for. These words, such as *(ḍaráb)tukum*, conform to the "foot extrametricality" pattern identified by Hayes (1995, 232). The most important competing candidate is **ḍarab(túkum),* which also satisfies NON-FINALITY but violates FT-FORM. In contrast, the actual output form *(ḍaráb)tukum* satisfies NON-FINALITY and FT-FORM at the price of inferior rightward alignment of its main stress (ALIGN-HEAD (McCarthy and Prince 1993a)). We therefore have the ranking NON-FINALITY >> FT-FORM >> ALIGN-HEAD, as the following tableau certifies:

(25) Iambic Stress in *(ḍaráb)tukum*

			NON-FINALITY	FT-FORM	ALIGN-HEAD
☞	(a)	(ḍaráb)tukum			**
	(b)	ḍarab(túkum)		*!	*
	(c)	ḍarab(tukúm)	*!		

Now we can get to the point: iambic prosody supplies 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 I/TL or GH, iambic feet optimally match their weak-strong prominence with short-long quantity. In more familiar cases, 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:[14]

---

[14]There are various precedents in the literature for positing moraless syllables. Hyman (1985) and Piggott(1995) argue that epenthetic vowels are weightless. Piggott also (p. 288) reviews some of the literature where degenerate syllables have been assumed under other circumstances. Indeed, degenerate syllables have frequently been proposed in analyses of Arabic (Aoun 1979; Broselow 1992; Farwaneh 1995; McCarthy and Prince 1990a; McCarthy and Prince 1990b; Selkirk 1981).

(26) The ΔL Iamb in Trisyllabic Deletion Cases

    a. ktíbat   b. ʔinksárat

     Ft   Ft

$\sigma_w \quad \sigma_s$ $\sigma_w \quad \sigma_s$

      μ μ
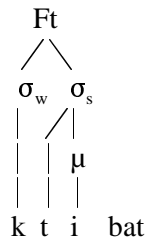
k t i   bat   ʔin k s a rat
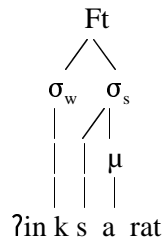
There are several reasons to think that this account of trisyllabic deletion is essentially correct.

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

(27)

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 I/TL or GH is satisfied without further ado: /katabna/ → *(ki.táb)na*, with a LH iamb.
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 (25): /rama/ → *(ríma)* 'he threw'.

Second, this analysis explains a significant correlation in the 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). Indeed, Kager's analysis of Macushi Carib is a close parallel in many respects.

Fourth, this analysis makes sense syllabically. Biconsonantal initial and triconsonantal medial clusters occur *only* as a result of vowel deletion, supporting the claim that they actually involve degenerate syllables : *(k.tí).bat*, *ʔin.(k.sá).rat*.[15] This too is closely paralleled in Kager's analysis of Macushi Carib.

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.

---

[15]Al-Mozainy (1981) reports syllabifications like *kti.bat* and *yak.tbin*, based on his own intuitions. Since surface degenerate syllables were not an available theoretical option at the time, he had no choice but to assign the extra consonant to one of the visible syllables. Presumably *yak.tbin* is analogized to *kti.bat*, where the choice is forced.

The details of the analysis are instructive, because they lead to issues about the opacity of phonological derivations. For concreteness, I adopt the following formulation of Grouping Harmony, after Prince (1990):[16]

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

To compel deletion of underlying /a/, GRPHARM must be ranked above IDENT-ADJ(Dur), which otherwise blocks /a/-deletion:

(29) Application to Trisyllabic Deletion Case

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

The failed candidate (29b) contains a LL iambic foot, violating GRPHARM. The alternative in (29a) contains an iamb of properly unequal weight, obtained by deleting the first vowel, leaving only a weightless syllable behind. Alternative candidates like *(ktí).bat (with a complex onset), *(tí).bat (with consonant deletion), or *(kitá:).bat (with lengthening instead of shortening) violate undominated constraints.

The DY case *(ʔaká)lat* is analyzed in much the same way, except that it shows the effect of an undominated constraint against a degenerate syllable with ʔ. Ranked above GRPHARM, that constraint effectively blocks trisyllabic deletion in words with initial ʔ.

(30) Application to /ʔakal‒at/

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

In other words, 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.

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 (29):

---

[16]A reminder: there are two different systems at play here ─ intrinsic or phonetic duration (relevant to REDUCE, IDENT(Dur)) and moraic or phonological quantity (relevant to GRPHARM). They do not appear to interact.

(31) Tableau (29) with *(*kát*).bat* Added

		/katab-at/	GRPHARM	IDENT-ADJ(Dur)	REDUCE	IDENT(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. 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.

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*. A speculative reformulation of GRPHARM would instead favor *(*s.mí). ʕat*. The following table makes this problem clear:

(32) 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* is.

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

(33) Opacity (Kiparsky 1973b) (neutralization case suppressed)
    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*.

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 (cf. clause (a) of the definition). In Optimality Theory, the hallmark of opacity is unexplained markedness or faithfulness violation by the actual output form (McCarthy 1998; to appear). In (31), as was already noted, the intended output *(k.tí).bat* has unexplained violations of both the markedness constraint REDUCE and the faithfulness constraint IDENT(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 (31) 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 1998; to appear),[17] 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 IDENT-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 ✯IDENT-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) and stress is iambic (though not perfectly in accord with GRPHARM).

The sympathy constraint, here designated by ❀SYM, evaluates candidates for similarity (in a way to be made precise below) to the sympathetic candidate. The actual output form *(k.tí).bat* (34a) is more similar to ❀*(ki.tí).bat* than **(kát).bat* (34b) 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 (31):

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

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

---

[17]Other work on sympathy theory includes Davis (1997a; 1997b), de Lacy (1998), Dinnsen et al. (1998), Ito and Mester (1997a; 1997b; 1998), Karvonen and Sherman (1997; 1998), Merchant (1997), Sanders (1997), and Walker (1998).

To complete the picture, we need to check that this application of sympathy has no untoward consequences for the rest of the language. No effects of sympathy are possible 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 (15) shows that deletion of /a/, which translates into violation of ✳IDENT-ADJ(Dur), only occurs in the /katab-at/ → *(k.tí).bat* derivation, and so that is the only circumstance where sympathy is relevant.

5. Sympathy and Cumulativity

The issue to be addressed now is the nature of the sympathy constraint ✾SYM. In 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 (34), 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 the inter-candidate faithfulness constraints of previous work on sympathy and in favor of a very different alternative, based on evaluating whether two candidates were produced by the same unfaithful mappings.

Perhaps the clearest example of the excessive descriptive power of inter-candidate faithfulness is the quasi-Yokuts example (13), repeated here for convenience:

(13) Quasi-Yokuts
	Underlying	/maat/	
⇨	Epenthesis	maati	To repair trimoraic syllable.
	Palatalization	maači	*ti* → *či* generally.
⇨	Apocope	maač	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 (35):

(35) Constraint Rankings for Quasi-Yokuts
a.	*$[\mu\mu\mu]_\sigma$ >> DEP-V	Trimoraic syllables are repairable by epenthesis.
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 visibly occurs is determined by the ranking in (35c), 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 ✾*maači*, 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 inter-candidate sympathetic faithfulness constraint ✿IDENT(high). By dominating its input-output counterpart IDENT(high), sympathetic ✿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:

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

	/maat/	*µµµ]$_\sigma$	*ti	✿IDENT(high)	IDENT(high)	DEP-V	✯MAX-µ
Transparent ☜ (a)	mat			*!			*
Sympathetic ✿ (b)	maači				*	*!	✔
Opaque ☞ (c)	mač				*		*
(d)	maat	*!		*!			✔
(e)	maati		*!	*!		*	✔
(f)	maač	*!			*		✔

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 ✿*maači* 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, and they need not detain us further.

This analysis pretty effectively simulates a feeding DY derivation. The input /maat/ is mapped onto the output *mač* through sympathetic attraction to ✿*maači*. On the assumption that such cases are actually impossible, we have to conclude that the theory of sympathy set forth in McCarthy (1998) is too powerful.

What's 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.

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.[18] 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 ✿*maači*'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, precisely specifying the type and locus of unfaithfulness. 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 special positions (Beckman 1997; Beckman 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. In specifying the locus of violation, we can reference elements of the input string by their indices. We can therefore talk about two distinct unfaithful mappings affecting /$p_1 a_2 p_3$/: 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.[19]

The locus of epenthesis is usually defined on the output; to keep things simple, it would be nice to have a way of talking about the locus of epenthesis relative to the input. One move, suggested to me by Alan Prince, is to 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 instance, input /a/ is equivalent to /$\varepsilon_{0\text{-}1} a_1 \varepsilon_{1\text{-}1} \varepsilon_{1\text{-}2}$/ which maps onto output $ʔ_{0\text{-}1} a_1 ʔ_{1\text{-}1} i_{1\text{-}2}$.

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. Indeed, syllabification aside, 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 the sympathetic inter-candidate faithfulness constraints of McCarthy (1998).

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

[18]I am greatly indebted to Alan Prince for discussion of this material.

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

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 *cumulative* with respect to the other.

- $U_{Cand1} \subsetneq U_{Cand2}$. In this case, Cand1 and Cand2 are different but comparable.[20] Cand2 is 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.

There is a somewhat more perspicuous way of looking at these inter-candidate relations. What we are talking about here is a partial ordering among candidates. The one 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 of McCarthy (1998). Underlying /deš?/ is mapped onto surface *deše* by two unfaithful mappings, epenthesis of *e* and deletion of *?*. Interesting candidates include sympathetic ❀*deše?* and the transparent competitor **deš*, where *?* has deleted without the seemingly superfluous epenthesis process.

(37) Partial Ordering Diagram for Hebrew /$d_1e_2\check{s}_3?_4$/

	deš?		i.e.,		Faithful	
✎deš		❀deše?		MAX@4		DEP@3-1
	☞deše				Both	
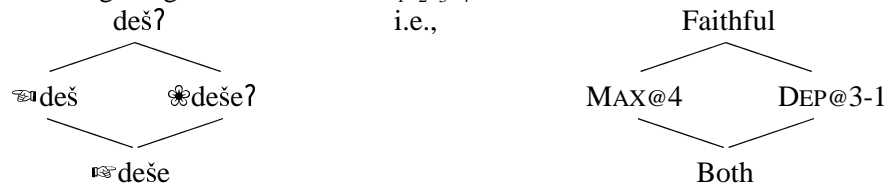
Obviously, this is just a tiny portion of the infinite set of candidates. Standing at the top of the partial ordering is the most faithful candidate (see Moreton (1996) on the existence of this candidate). Below it, on the first tier, are candidates with a single unfaithful mapping, including deletion of /?/ 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, *deš?*, is comparable with all other candidates, and all other candidates accumulate its unfaithful mappings. (That's 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 ❀*deše?*. Significantly, *deše*'s transparent competitor **deš* is not cumulative with respect to the sympathetic candidate ❀*deše?*. This non-cumulativity proves to be fatal.

A similar diagram can be constructed for another of the examples in McCarthy (1998), the failure

---

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

of raising to occur in Bedouin Arabic words like *ba.du*, derived by glide vocalization from underlying /badw/:

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

<div align="center">

❀badw.   i.e.,   Faithful

bidw.   ☞ba.du   IDENT(high)@2   *DEP-μ@4

☜bi.du   Both

</div>
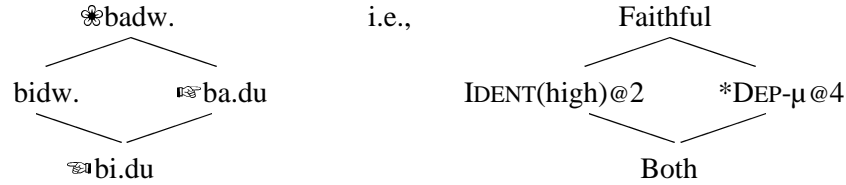
The sympathetic candidate is one in which the underlying glide has not vocalized, and so raising, which occurs in an open syllable, is not motivated. In this case, then, the sympathetic candidate is faithfully mapped from the input. The actual output form has glide vocalization but not raising; it competes 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.

This discussion gives a pretty good idea of how ❀SYM must evaluate candidates for similarity to the sympathetic candidate. In (37), it is crucial that ❀SYM judge *deše* to be more harmonic than its transparent competitor *deš, while in (38), ❀SYM must select *ba.du* over **bi.du*.[21] I offer the following definition as a first approximation:[22]

(39) ❀SYM

Given a sympathetic candidate ❀-Cand, to evaluate a candidate E-Cand, derived from the same input:

- If $U_{❀\text{-Cand}} \subseteq U_{\text{E-Cand}}$ then E-Cand's performance on ❀SYM is proportional to the cardinality of the set $U_{\text{E-Cand}} - U_{❀\text{-Cand}}$.
- If E-Cand and ❀-Cand are *non-comparable* in their unfaithful mappings, then E-Cand's performance on ❀SYM is worse than that of any candidate that is comparable.

(The notation "X−Y" denotes the relative complement of Y with respect to X. It is the set of elements in X that are not also in Y.) If E-Cand accumulates ❀-Cand's unfaithful mappings, then the distance between them, measured by the number of unshared unfaithful mappings, is a measure of their dissimilarity. In the limiting case, where ❀SYM is evaluating the sympathetic candidate itself, that distance is zero. On the other hand, if E-Cand and ❀-Cand are non-comparable, the distance between the two candidates is greater than the distance between any two comparable candidates. In this case, ❀SYM acts as a kind of filter, extinguishing all candidates which are not cumulative with respect to ❀-Cand. This special situation will be indicated with the symbol "⊥", for bottom.[23]

---

[21]It doesn't matter how the other candidates in (37) and (38), including the sympathetic candidates themselves, fare on ❀SYM, because they are satisfactorily disposed of by undominated markedness constraints like *COMPLEX.

[22]This definition is intentionally incomplete, omitting the case $U_{\text{E-Cand}} \supset U_{❀\text{-Cand}}$ and saying nothing about systems with multiple sympathetic candidates. Consult McCarthy (to appear) for further refinements.

[23]Alan Prince suggests an alternative path for technical development. The definition of ❀SYM in (39) could be decomposed into a fixed hierarchy of two constraints. The higher-ranking one, ❀CUMUL, tests categorically for cumulativity of the ❀-Cand/E-Cand relation. The lower-ranking constraint, ❀DIST, gradiently assesses the unshared

Here's a shortcut that uses a diagram like (37) or (38). If there is a purely downward path from ✿-Cand to E-Cand, then the number of links in that path is the number of marks on ✿SYM that E-Cand incurs. If there is no purely downward or purely upward path from ✿-Cand to E-Cand, then E-Cand incurs the definitively bad mark "⊥". Intuitively, the force of ✿SYM is to winnow the set of viable candidates. Only those candidates that are in the down-set[24] of ✿-Cand will survive, and in that down-set it will favor those that are closest to ✿-Cand.

Applied to (37) and (38), the definition of ✿SYM in (39) gives exactly the right results. In (37), ✿SYM correctly favors *deše* over its transparent competitor **deš*, relative to the sympathetic candidate ✿*deše7*. Because *deše* is cumulative with respect to ✿*deše7*, the distance between them is determined by the size of the set of unshared unfaithful mappings, or one "*". But **deš* is not cumulative with respect to ✿*deše7*, so **deš*'s performance is worse than that of any candidate that is comparable to ✿*deše7*. As required, then, evaluation by ✿SYM selects *deše* over **deš*. Likewise, in (38), ✿SYM judges *ba.du* to be more harmonic than its transparent competitor **bi.du*. 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 "*" to **bi.du*'s two "*"s.

This definition of ✿SYM also successfully contends with the more complex example that brought us here, the opacity of *(k.tí).bat* in (34). It's necessary for ✿SYM to favor *(k.tí).bat* over transparent **(kát).bat* relative to the sympathetic candidate ✿*(ki.tí).bat*. The definition of ✿SYM in (39) does exactly that. Consider the sets of unfaithful mappings associated with these three candidates:[25]

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

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

Except for addition of the indices that mark the locus of faithfulness violation, these values for the unfaithful mappings are exactly what can be read off the tableau (34).

The definition of ✿SYM in (39) tells us how to evaluate the other two candidates for similarity with ✿*(ki.tí).bat*. The intended output form *(k.tí).bat* has a proper superset of ✿*(ki.tí).bat*'s faithfulness violations. The cardinality of $U_{(k.tí).bat} – U_{(ki.tí).bat}$ is 1, so *(k.tí).bat* receives one "*" from ✿SYM. (As usual, the sympathetic candidate itself performs perfectly on ✿SYM.) But the transparent competitor **(kát).bat* has a disjoint set of unfaithful mappings from the ✿-candidate's. They are non-comparable or, equivalently, non-cumulative, and so **(kát).bat*'s performance on ✿SYM is worse than *(k.tí).bat*'s, here indicated by "⊥". The following tableau updates (34) to reflect these developments:

---

unfaithful mappings.

[24]A *down-set* is a subset Q of a poset P such that if $x \in Q$ then so is every $y \in P$ for $y \leq x$.

[25]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 1995; Alderete 1998; Bye 1996; Inkelas 1999; Ito, Kitagawa, and Mester 1996; McCarthy 1995; Pater 1995). Violations of IDENT(Stress), although they do not affect the outcome in in (40), are part of the package of unfaithful mappings that each candidate brings with it.

(41) ❋SYM Applied to /katab-at/ → *(k.tí).bat*, **(kát).bat*

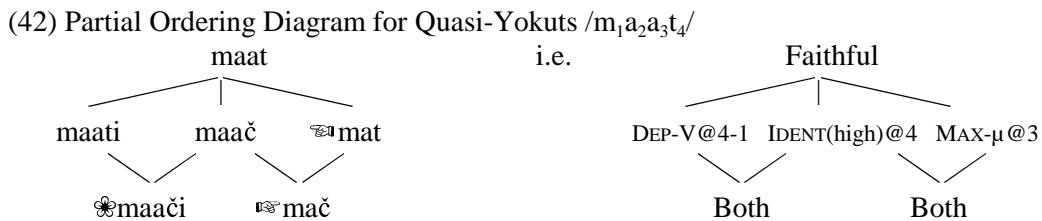|		/katab-at/	GRPHARM	❋SYM	✹IDENT-ADJ(Dur)	REDUCE	IDENT(Dur)
Opaque	☞	(a) (k.tí).bat		*	*	*	**
Sympathetic	❀	(b) (ki.tí).bat	*!		✓	**	**
Transparent	☜	(c) (kát).bat		⊥ !	*		*

Observe that ❋SYM is crucially dominated by GRPHARM. This reflects a general property of sympathy under the new regime: since the sympathetic candidate performs perfectly on ❋SYM, ❋SYM must be crucially dominated if the sympathy effect is to be non-vacuous.

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 presented a more restrictive alternative, in which the only information that can be transmitted from the sympathetic candidate is the set of its unfaithful mappings. This information is conveyed by the constraint ❋SYM, which ranks candidates according to how well they match the sympathetic candidate in unfaithful mappings. The most harmonic candidate under the ❋SYM regime is the sympathetic candidate itself; next most harmonic is a candidate which has all of the sympathetic candidate's unfaithful mappings, plus one more; and so on. ❋SYM maximally disdains candidates whose unfaithful mappings are not cumulative with respect to those of the sympathetic candidate.

This approach is very different from the system of inter-candidate faithfulness constraints developed in McCarthy (1998). ❋SYM measures difference between candidates indirectly, in terms of input-output faithfulness, rather than directly, as in the original sympathy implementation.

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 (36). 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 (cf. (13)). 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.

To see this concretely, consider the following diagram, which organizes the quasi-Yokuts candidates in (36) according to their unfaithful mappings:

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

```
        maat                    i.e.            Faithful
      /   |   \                              /      |      \
  maati  maač  ☜mat              DEP-V@4-1  IDENT(high)@4  MAX-μ@3
        \   \    \                              \      |      /
      ❀maači  ☞mač                          Both          Both
```

It is immediately evident that the intended output *mač* does not accumulate the unfaithful mappings of the sympathetic candidate ❀*maači*. ❋SYM extinguishes all non-cumulative candidates, leaving only ❀*maači*'s down-set as viable. Thus, ❋SYM is equally fatal to *mač* and its transparent competitor *mat,

assigning both the worst possible violation-marks. Because ❀SYM cannot do the job that would be required of it, 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 Syllabic Structure

❀SYM and the related notion of cumulativity are defined in terms of shared unfaithful mappings. Indeed, the candidates themselves can be described in terms of the unfaithful mappings that produce 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 in UG, and so its unexpected presence or absence will count as an unfaithful mapping for the purposes of ❀SYM. 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 whether a derivation is cumulative or how well a candidate performs on ❀SYM.

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*. 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:

(43) Faithfulness-Free Syllabification
      No constraints of UG demand faithfulness to syllables *per se*.

This section explores the implications of (43) 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 prohibited feeding DY derivations, but they will involve non-monotonicity *only* in faithfulness-free syllabification.[26]

The example back in (3) is a hypothetical, though undoubtedly authentic-appearing, case of this sort.

---

[26]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*.

A genuine example 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-liinda* 'to wait', *mu-leenzi* 'boy', *muu-ntu* 'person', *baa-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 in Clements's analysis:

(44) Luganda Derivation with Resyllabification

$$
\begin{array}{ll}
& \quad\quad\quad \mu\;\;\mu \\
& \quad\quad\quad |\;\;\; | \\
\Rightarrow \text{Syllabification} & \quad\quad \text{mun.tu} \\[4pt]
& \quad\quad\quad \mu\mu\;\;\mu \\
& \quad\quad\quad ||\;\;\; | \\
\text{Weight by Position} & \quad\quad \text{mun.tu} \\[4pt]
& \quad\quad\quad \mu\mu\;\;\mu \\
& \quad\quad\quad |\;\;\;\;\; | \\
\Rightarrow \text{Prenasalization} & \quad\quad \text{mu .}^{n}\text{tu} \\[4pt]
& \quad\quad\quad \mu\mu\;\;\mu \\
& \quad\quad\quad |/\;\;\;\; | \\
\text{Spreading} & \quad\quad \text{mu .}^{n}\text{tu}
\end{array}
$$

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 processes assigning positional weight and creating prenasalized consonants.[27]

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:

(45) NO-CODA, MAX >> UNIFORMITY

/muntu/	NO-CODA	MAX	UNIFORMITY
☞ (a) muu.^{n}tu			*
(b) mun.tu	* !		
(c) muu.tu		* !	

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

This analysis is not sufficient, however, because there is an element of opacity in Luganda

[27]For a comrephensive analysis of Luganda within OT, see Rosenthall (1994).

coalescence, as I have noted. The mapping /muntu/ → *muu.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 underlining marks the *n* as moraic, as in the second step of (44)). The selector constraint is ☆UNIFORMITY, which is obeyed by ✿*mun.tu* and violated by the actual output form *muu.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. Sympathy must favor opaque *muu.ⁿtu* over transparent **mu.ⁿtu* 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:

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

	Candidate	U$_{Candidate}$
☞	muu.ⁿtu	DEP-μ@1-1,[28] UNIFORMITY@3,4
✿	mun.tu	DEP-μ@1-1
✑	mu.ⁿtu	UNIFORMITY@3,4

The intended output form has a superset of the sympathetic candidate's unfaithful mappings; they are in a relationship of cumulativity. But the transparent competitor does not accumulate the sympathetic candidate's unfaithful mappings. Therefore, ✿SYM will favor *muu.ⁿtu* over **mu.ⁿtu*, exactly as desired. The following tableau completes the argument at the level of formal detail:

(47) Luganda /muntu/ → *muu.ⁿtu*

		/muntu/	MAX	WxP	NO-CODA	✿SYM	☆UNIF	DEP-μ
☞	(a)	muu.ⁿtu				*	*	*
✿	(b)	mun.tu			* !		✓	*
✑	(c)	mu.ⁿtu				⊥ !	*	
	(d)	muu.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 ✿SYM, the sympathetic candidate bestows its favor on *muu.ⁿtu*, which accumulates its unfaithful mappings, over **mu.ⁿtu*, which does not. Since the sympathetic candidate itself is guaranteed perfect performance on ✿SYM, its worst violation, NO-CODA, must be ranked above ✿SYM.[29]

---

[28] I assume that epenthesis on the moraic tier is treated like epenthesis on the segmental tier.

[29] With some adjustments, this analysis of Luganda can be adapted to the Ponapean example discussed by Davis (1997a; 1997b)

This analysis succeeds under the assumptions that (i) ✽SYM and cumulativity are defined in terms of shared unfaithful mappings and (ii) syllabification is not an unfaithful mapping — i.e., (43). 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:

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

	Candidate	$U_{Candidate}$
☞	muu.^{n}tu	DEP-μ@1-1, UNIFORMITY@3,4
		$m_1 \rightarrow$ Onset, $u_2 \rightarrow$ Nucleus, $n_3 \rightarrow$ Onset, $t_4 \rightarrow$ Onset, $u_5 \rightarrow$ Nucleus
✽	mu<u>n</u>.tu	DEP-μ@1-1
		$m_1 \rightarrow$ Onset, $u_2 \rightarrow$ Nucleus, $n_3 \rightarrow$ Coda, $t_4 \rightarrow$ Onset, $u_5 \rightarrow$ Nucleus
☜	mu.^{n}tu	UNIFORMITY@3,4
		$m_1 \rightarrow$ Onset, $u_2 \rightarrow$ Nucleus, $n_3 \rightarrow$ Onset, $t_4 \rightarrow$ Onset, $u_5 \rightarrow$ Nucleus

Observe that the sympathetic candidate has $n_3 \rightarrow$ Coda, while the intended output has $n_3 \rightarrow$ Onset. Thus, there is no cumulativity relation between these candidates. This means that the intended output ties on ✽SYM performance with its transparent competitor *mu.^{n}tu*. (Both are non-cumulative, so both receive the fatal "⊥" mark.) This tie is disastrous; inspection of (47) shows that a tie between *muu.^{n}tu* and *mu.^{n}tu* on ✽SYM leaves the decision up to DEP-μ, which *muu.^{n}tu* then fatally violates. That's 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, exploring their ramifications and showing how they are compatible with the thesis in (43) that syllabification is not regulated by faithfulness constraints.

+++++++++++++++++++

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 may demand 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 the usual assumptions about 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 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 1990a; McCarthy and Prince 1986; 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.[30]

<div align="center">+++++++++++++++++++</div>

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; Borgstrøm 1940; Bosch 1991; Bosch and de Jong 1997; Bosch and de Jong 1998; Clements 1986b; 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 minimal pairs differing in syllabification, like *šeR.ak* 'a glass of whiskey' (from /šæRak/) and *šæ.rak* 'to fade' (from /šærk/).

Following Clements (1986b), Blevins (1995, 231) sketches a plausible derivational analysis:

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

Underlying	/šæRak/	/šærk/
Syllabification	šæ.Rak	šæ.rk
Attraction	šæ<u>R</u>.ak	—
Epenthesis	—	šæ.rak

The source of the surface contrast is the counter-feeding order between Attraction and Epenthesis. Attraction makes the initial syllable, which is stressed, 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.

---

[30]Establishing sufficient conditions to rule out syllable-copying reduplication is a difficult problem which I will not try to address here.

Clearly, Barra Gaelic does not require syllabic faithfulness in the input→output mapping, so it presents no difficulties for my main premise. Furthermore, the opaque interaction seen in (49) is compatible with the theory of sympathy. Sympathy is needed to explain why underlying /šærk/ maps onto surface *šæ.rak* instead of its competitor **šær.ak* (which transparently shows the effect of the constraint responsible for attraction). The obvious selection for sympathetic candidate is fully faithful ✿*šæ.rk*, which is the most harmonic candidate that obeys ☆DEP-μ. Here, then, are the unfaithful mappings for the relevant candidates:

(50) Unfaithful Mappings Relative to Input /š$_1$æ$_2$r$_3$k$_4$/

Candidate	/šærk/U$_{Candidate}$
☞ šæ.rak	DEP@3-1, DEP-μ@3-1
✿ šæ.rk	Ø
☜ šær̲.ak	DEP-μ@3, DEP@3-1, DEP-μ@3-1

Vowel epenthesis incurs DEP violations for both the vowel itself and the added mora associated with that vowel. In addition, **šær̲.ak* has a DEP-μ violation for the mora added to the *r* to make the first syllable heavy.

According to (50), both the intended output *šæ.rak* and its transparent competitor **šær̲.ak* stand in a cumulativity relation with the sympathetic candidate ✿*šæ.rk*. This situation is equivalent to the /badw/ example in (38): ✿SYM favors the candidate which is closer, in terms of shared unfaithful mappings. That candidate is *šæ.rak*, since **šær̲.ak* has an additional DEP-μ violation. By ranking ✿SYM above the constraint responsible for the attraction effect, the correct outcome is assured. It seems clear, then, that Barra Gaelic is fully compatible with the theory presented here, and may even provide positive support for it.[31]

<center>+++++++++++++++++++</center>

Contrasts in quantity and syllabicity. It is widely accepted that contrasts of quantity and syllabicity are represented by deploying moras in underlying representation (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 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 a Surrogate for Syllabic Faithfulness
    a. /pata/   →   *pa.ta*
    b. /pat̲a/   →   *pat̲.a*

A language with mappings like these would be one in which faithfulness to moras in underlying

---

[31]Barra Gaelic may also be simply irrelevant — see Bosch and de Jong (1997; 1998) and Beckman (1998), where the relevant difference is attributed to suprasyllabic structure (stress or a "supersyllable" constituent).

representation produces contrast in syllables in surface representation. To account for the impossibility of tautomorphemic syllabic contrast, it's necessary (though not sufficient) to universally rule out the system with these mappings.

In OT, 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's *always* a more harmonic mapping, under *any* permutation of the constraints of UG. Suppose that UG consists of only the following constraints:

(52) A UG Constraint Set for Non-Contrastive Syllabification
    a. ONSET
            $*[_\sigma V$
    b. NO-CODA
            $*C]_\sigma$
    c. $*\mu_{Cons}$ (Sherer 1994, 26)
            $*[C]_\mu$ (Consonants may not be parsed as moraic.)
    d. WxP
            If $C_i]_\sigma$, 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 /pat̪a/ onto just *pat̪.ta* or *pa.ta*, depending on the disposition of MAX-μ relative to the structural constraints. No permutation will produce *pat̪.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.[32]

Other theories of UG may have other consequences. In particular, if UG contains a constraint that specifically militates against geminate 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:[33]

---

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

[33]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/ → *pat̪ka* 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/ → *pat̪ka* 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).

(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/ → *pat̪.a*.

<div align="center">+++++++++++++++++++</div>

To summarize the results of this section, I have argued that there are no constraints enforcing faithfulness to syllables per se. An endo-theoretic argument, based on applying revised sympathy to cases like Luganda, was supported by exo-theoretic 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, does not necessarily 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 undoubtedly 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 been little studied in the intervening decades; it is certainly significant that these questions still claim our attention.

References

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

Alderete, John. 1995. Faithfulness to prosodic heads. Ms., University of Massachusetts at Amherst, Amherst, MA.

Alderete, John. 1998. Morphologically-governed accent in Optimality Theory. Doctoral dissertation, University of Massachusetts, Amherst, Amherst, MA.

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.

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.

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

Beckman, Jill N. 1998. Positional Faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.

Benua, Laura. 1997a. Affixes Classes are defined by Faithfulness. In *University of Maryland Working Papers in Linguistics 5. Selected Phonology Papers from Hopkins Optimality Theory Workshop 1997 / University of Maryland Mayfest 1997*, edited by V. Miglio and B. Moreen.

Benua, Laura. 1997b. Transderivational Identity: Phonological Relations between Words. Doctoral dissertation, University of Massachusetts, Amherst.

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.

Blevins, Juliette. 1995. The syllable in phonological theory. In *The Handbook of Phonological Theory*, edited by J. A. Goldsmith. Cambridge, MA, and Oxford, UK: Blackwell.

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

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 *Studies in Lexical Phonology*, edited by S. Hargus and E. Kaisse. New York: Academic Press.

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

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

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.

Broselow, Ellen. 1992. Parametric variation in Arabic dialect phonology. In *Perspectives on Arabic Linguistics*, edited by E. Broselow, M. Eid and J. McCarthy. Amsterdam and Philadelphia: John Benjamins.

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.

Bye, Patrik. 1999. Coherence in Finnish consonant gradation: Towards a theory of phonological lenition. In *CLS 32*. Chicago, IL: Chicago Linguistic Society.

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

Causley, Trisha. 1997. Identity and featural correspondence: The Athapaskan case. *NELS* 27:93-105.

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 *Studies in Compensatory Lengthening*, edited by L. Wetzels and E. Sezer. Dordrecht: Foris.

Clements, G. N. 1986b. Syllabification and epenthesis in the Barra dialect of Gaelic. In *The Phonological Representation of Suprasegmentals*, edited by K. Bogers, H. v. d. Hulst and M. Mous. Dordrecht: Foris.

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

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

Davis, Stuart. 1997a. The flowering of Optimality Theory: Ponapean nasal substitution and the problem of intermediate forms. Ms., Indiana University, Bloomington, IN.

Davis, Stuart. 1997b. A sympathetic account of nasal substitution in Ponapean. *Phonology at Santa Cruz* 5:15-28.

de Lacy, Paul. 1998. Sympathetic stress. Ms., University of Massachusetts, Amherst, Amherst, MA.

Dinnsen, Daniel A., Laura W. McGarrity, Kathleen O'Connor, and Kim Swanson. 1998. On the role of sympathy in acquisition. Ms., Indiana University, Bloomington, IN.

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

Gafos, Adamantios, and Linda Lombardi. In preparation. Coronal transparency. Ms., New York University and University of Maryland, College Park, New York, NY and College Park, MD.

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.

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

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 *Derivations and Constraints in Phonology*, edited by I. Roca. Oxford: Clarendon Press.

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

Harris, John. 1990. Derived phonological contrasts. In *Studies in the Pronunciation of English*, edited by S. Ramsaran. London: Routledge.

Harris, James. 1993. Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by K. Hale and S. J. Keyser. Cambridge, MA: MIT Press.

Hastings, Ashley. 1974. Stifling. Ms., Indiana University Linguistics Club, Bloomington, IN.

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

Hayes, Bruce. 1987. A revised parametric metrical theory. In *Proceedings of NELS 17*, edited by J. McDonough and B. Plunkett. Amherst: GLSA.

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.

Hulst, Harry van der, and Jeroen van de Weijer. 1995. Vowel harmony. In *The Handbook of Phonological Theory*, edited by J. A. Goldsmith. Cambridge, MA, and Oxford, UK: Blackwell.

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 *The Prosody-Morphology Interface*, edited by R. Kager, H. v. d. Hulst and W. Zonneveld. Cambridge: Cambridge University Press.

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.

Ito, Junko, and Armin Mester. 1997a. Featural sympathy: feeding and counterfeeding interactions in Japanese. *Phonology at Santa Cruz* 5:29-36.

Ito, Junko, and Armin Mester. 1997b. Sympathy theory and German truncations. In *University of Maryland Working Papers in Linguistics 5. Selected Phonology Papers from Hopkins Optimality Theory Workshop 1997 / University of Maryland Mayfest 1997*, edited by V. Miglio and B. Moren.

Ito, Junko, and Armin Mester. 1998. German coda conditions: Sympathy meets freedom of the input. Paper presented at The Syllable: Typology and Theory, Tuebingen, Germany.

Iverson, Gregory K. 1995. Rule ordering. In *The Handbook of Phonological Theory*, edited by J. A. Goldsmith. Cambridge, MA, and Oxford, UK: Blackwell.

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 *Derivations and Constraints in Phonology*, edited by I. Roca. Oxford: Oxford University Press.

Karvonen, Daniel, and Adam Sherman. 1997. Sympathy, opacity, and u-umlaut in Icelandic. *Phonology at Santa Cruz* 5:37-48.

Karvonen, Dan, and Adam Sherman. 1998. Opacity in Icelandic revisited: a Sympathy account. In *Proceedings of NELS 28*.

Keer, Edward. 1999. Geminates, the OCP, and faithfulness. Doctoral dissertation, Rutgersity University, New Brunswick, NJ.

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

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 *A Festschrift for Morris Halle*, edited by S. Anderson and P. Kiparsky. New York: Holt, Rinehart and Winston.

Kiparsky, Paul. 1973b. Phonological representations. In *Three Dimensions of Linguistic Theory*, edited by O. Fujimura. Tokyo: TEC.

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

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

Lamontagne, Greg, and Keren Rice. 1995. A correspondence account of coalescence. In *University of Massachusetts Occasional Papers in Linguistics 18*, edited by J. Beckman, S. Urbanczyk and L. Walsh Dickey. Amherst, MA: GLSA.

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.

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

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

McCarthy, John. 1991. Synchronic Rule Inversion. In *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Soceity*, edited by L. Sutton, C. Johnson and R. Shields. Berkeley, CA: Berkeley Linguistics Society.

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

McCarthy, John. 1993b. Containment, consistency, and alignment. Paper presented at Rutgers Optimality Workshop I, New Brunswick, NJ.

McCarthy, John. 1994. On coronal 'transparency'. Paper presented at TREND, Santa Cruz, CA.

McCarthy, John. 1998. Sympathy and phonological opacity. Ms., University of Massachusetts, Amherst, Amherst, MA.

McCarthy, John. to appear. Sympathy and phonological opacity. *Phonology*.

McCarthy, John, and Alan Prince. 1988. Quantitative transfer in reduplicative and templatic morphology. In *Linguistics in the Morning Calm*, edited by Linguistic Society of Korea. Seoul: Hanshin Publishing Co.

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 *Perspectives on Arabic linguistics II: Papers from the Second Annual Symposium on Arabic Linguistics*, edited by M. Eid and J. McCarthy. Amsterdam: John Benjamins.

McCarthy, John, and Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *University of Massachusetts Occasional Papers in Linguistics UMOP 18*, edited by J. Beckman, S. Urbanczyk and L. Walsh Dickey. Amherst, MA: GLSA.

McCarthy, John J. 1995. Extensions of faithfulness: Rotuman revisited. Ms., University of Massachusetts, Amherst, .

McCarthy, John J., and Alan S. Prince. 1986. Prosodic Morphology. Ms., University of Massachusetts at Amherst and Brandeis University, .

McCarthy, John J, and Alan S. Prince. 1993a. Generalized Alignment. In *Yearbook of Morphology*, edited by G. Booij and J. v. Marle.

McCarthy, John J., and Alan S. Prince. 1993b. Prosodic Morphology I: Constraint Interaction and Satisfaction. Ms., University of Massachusetts, Amherst, and Rutgers University, .

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* 23:369-89.

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

Moravcsik, Edith. 1978. Reduplicative constructions. In *Universals of Human Language*, edited by J.

Greenberg. Stanford: Stanford University Press.

Moreton, Elliott. 1996. Non-computable functions in Optimality Theory. Ms., University of Massachusetts, Amherst, Amherst, MA.

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. Ms., UC Santa Cruz, Santa Cruz, CA.

Pater, Joe. 1995. On the nonuniformity of weight-to-stress and stress preservation effects in English. Ms., McGill University, .

Pater, Joe. 1999. Austronesian nasal substitution and other *NC effects. In *The Prosody-Morphology Interface*, edited by R. Kager, H. v. d. Hulst and W. Zonneveld. Cambridge: Cambridge University Press.

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

Prince, Alan. 1975. The phonology and morphology of Tiberian Hebrew. Doctoral dissertation, MIT, Cambridge, MA.

Prince, Alan. 1990. Quantitative consequences of rhythmic organization. In *Parasession on the Syllable in Phonetics and Phonology*, edited by M. Ziolkowski, M. Noske and K. Deaton. Chicago: Chicago Linguistic Society.

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

Prince, Alan. 1997. Elsewhere & otherwise. Ms., Rutgers University, .

Prince, Alan. 1998. Two lectures on Optimality Theory. Paper presented at Phonology Forum 1998, Kobe University.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, NJ.

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

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

Rosenthall, Samuel. 1994. Vowel/glide alternation in a theory of constraint interaction. Doctoral dissertation, University of Massachusetts, Amherst, Amherst, MA.

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

Sanders, Nathan. 1997. On sympathetic correspondence. *Phonology at Santa Cruz* 5:91-102.

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

Selkirk, Elisabeth. 1981. Epenthesis and degenerate syllables in Cairene Arabic. In *Theoretical Issues in the Grammar of the Semitic Languages*, edited by H. Borer and J. Aoun. Cambridge, MA: MIT.

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. Paper presented at talk, University of Arizona.

Walker, Rachel. 1998. Nasalization, neutral segments, and opacity effects. Doctoral dissertation, University of California, Santa Cruz, Santa Cruz, CA.

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

Zoll, Cheryl. 1993. Ghost segments and Optimality. In *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, edited by E. Duncan, M. Hart and P. Spaelti. Stanford, CA: Stanford Linguistics Association.