Limits on global rules in Optimality Theory with Candidate Chains [*]

Matthew Wolf
Yale University
matthew.wolf@yale.edu

Abstract: In OT with Candidate Chains (McCarthy 2007), candidates are multi-step derivations, and the PREC constraints which regulate the order of derivational steps can inspect entire candidate derivations. This means (Wilson 2006; Wolf 2008, 2010) that OT-CC opens the door to certain kinds of 'global rules' (Lakoff 1970)—that is, effects in which the application or non-application of a process is decided with crucial reference to derivational history. This paper investigates what limits may exist on OT-CC's global-rule powers, focusing on two forms of opacity which are possible under a theory where all rules apply simultaneously, but not under sequential rule-application: mutual counterfeeding and mutual counterbleeding. It is shown that the original version of OT-CC allows neither, but that each of them could be made possible with relatively simple revisions to the original theory. Possible examples of these forms of opacity are discussed.

[*] This paper has benefitted greatly from the comments, ideas, and questions I've received from Stephen R. Anderson, Maria Gouskova, Karen Jesney, Darya Kavitskaya, John McCarthy, E-Ching Ng, Joe Pater, Peter Staroverov, the editors, an associate editor, and two anonymous reviewers. Audiences at Rutgers University, particularly Mark Baker, Paul de Lacy, Paula Houghton, Shigeto Kawahara, Alan Prince, and Bruce Tesar, and at the 2010 LSA Meeting in Baltimore, particularly Eugene Buckley, Luigi Burzio, Bruce Hayes, Paul Kiparsky, and Jason Riggle, supplied helpful feedback on related material. Bill Foley and Gillian Gallagher kindly answered my queries about possible examples of global rules. I am grateful to all of them. The views presented here, as well as all of the mistakes, are my own.

1. Introduction

In rule-based theories of phonology, different positions can be and have been taken about the order in which phonological rules apply relative to one another. One possible position is that all of the rules which mediate the mapping from one representation onto another apply simultaneously. The principle of Local Determinacy, which was widely subscribed to in structuralist phonemics, states that the allophonic realization of each phoneme of an utterance can be determined solely with reference to the phonemic level itself. This is equivalent to assuming that allophonic rules all apply simultaneously (Anderson 1974: 29). If allophonic rules could apply one at a time, then the application of each rule would create a new representation intermediate between the phonemic level and the phonetic level, and it would then be possible for the allophonic realization of some phoneme to be crucially defined with respect to conditions on this intermediate level. The idea was also advanced that the rules that mapped morphophonemic representations onto phonemic ones also all applied simultaneously; Chomsky & Halle (1968: 19, fn. 5) credit Harris (1951: appendix to 14.32) with having 'first made explicit' this view.[1] Subsequent works exploring the possibility of simultaneous application include Chafe (1968), Ballard (1971), and Hyman (1993: §7.2).

Another position, probably more familiar nowadays, is the one staked out in the early years of generative phonology: that rules are ordered and apply one at a time.[2] During the 1960s and 70s when rule ordering was a matter of intense debate, two main arguments appeared against the position that rules all applied at the same time. The more widely-issued of these (Chomsky & Halle 1968: 349; McCawley 1968: 21-23; Postal 1968: chapter 7; Pullum 1976: 228-232) is that, if all rules apply simultaneously, feeding interactions are impossible, since no rule can apply to the output of any other rule, forcing the analyst in many cases to posit several rules which redundantly help enforce the same generalizations. Similar considerations arise with bleeding interactions (Bromberger & Halle 1989: 59-60).

The other argument which was made against all-simultaneous rule application is that it permits modes of rule interaction which (it was argued) are not attested, of which two main kinds have been noted. The first (Chomsky & Halle 1968: 19, fn. 5) is of a type that we may call *mutual counterfeeding*. In a mutual counterfeeding scenario, there are two rules, each of which creates strings which satisfy the structural description of the other rule, but *neither* rule applies to the output of the other. For

[1] On the other hand, a number of works in structuralist phonology did make use of ordered rules, notably Bloomfield (1939) and Wells (1949). For historical discussion, and comparison with proposals about rule ordering in generative phonology, see Kenstowicz (1976) and Goldsmith (2008).

[2] These are not, of course, the only two possible positions. Koutsoudas, Sanders & Noll (1974) and Pullum (1976), who reject the notion of extrinsic ordering, propose frameworks in which there are multi-step derivations, but where in the mapping from one step to the next, it is possible for multiple rules to apply simultaneously. Another notable theory is that of Local Ordering (Anderson 1974), in which rules apply one at a time, and may be extrinsically ordered, but in which the pairwise ordering of certain rules may be left unspecified, permitting universal ordering preferences to assign potentially different orderings to the rules, depending on the nature of the representation being operated on.

example, suppose that we have a language with the following two rules (both of which are eminently known to occur in real languages)[3]:

(1) /ə/→Ø / {V,#}(C)_(C){V,#}
(schwa deletes, except when a cluster of more than two consonants would result)

(2) /h/→Ø / _{[−voc],#}
(/h/ deletes before a consonant or glide, or word-finally, i.e. in coda position)

Each of these rules has the potential to feed the other. If syncope deletes a /ə/ which falls between an /h/ and another consonant, it will place the /h/ in preconsonantal position and make it eligible for /h/-deletion. Likewise, deleting an /h/ from a sequence /VhCəCV/ or /VCəhCV/ leaves the schwa with only one consonant on either side, making it eligible for syncope.

If these rules apply one at a time, there are two possible orders. If /ə/-syncope happens first, then /ə/-syncope will feed /h/-deletion, and /h/-deletion will counterfeed /ə/-syncope:

(3)

	/ehtəmu/	/ahəpi/
/ə/-syncope	*doesn't apply*	ahpi
/h/-deletion	etəmu	api

If /h/-deletion happens first, then /h/-deletion will feed /ə/-syncope, and /ə/-syncope will counterfeed /h/-deletion:

(4)

	/ehtəmu/	/ahəpi/
/h/-deletion	etəmu	*doesn't apply*
/ə/-syncope	etmu	ahpi

But suppose instead that it were possible for a grammar to specify that these two rules applied simultaneously, and that neither rule could re-apply later. This results in each of the two rules counterfeeding the other:

[3] Syncope rules which are blocked when they would create triconsonantal clusters are found in Hindi-Urdu (see discussion in §3.2 below) and in Yowlumne (Kisseberth 1970), among other languages. Deletion of coda/non-pre-vocalic /h/ occurs stem-finally in Hungarian (Vago 1977: 35) and also corresponds to a restriction in the static phonotactics of English, where [h] is banned in coda position.

(5)

/ehtəmu/→[etəmu]
(UR meets structural description of /h/-deletion but not /ə/-syncope)

/ahəpi/→[ahpi]
(UR meets structural description of /ə/-syncope but not /h/-deletion)

A second type of hypothetical interaction which is predicted under a theory with all-simultaneous application (Vago 1977; Baković 2007b) is one which we can call *mutual counterbleeding*. Here, there are two rules which are such that applying either rule to a form will cause the form to no longer meet the structural description of the other rule. However, when the input form meets the structural description of both rules, both rules nevertheless apply. Suppose, for instance, that we have a language with the same /h/-deletion rule as the previous example, along with a rule that vocalizes glides not adjacent to a vowel (as in Bedouin Arabic [McCarthy 1999 and references therein], or Hungarian [Vago 1977: 32]):

(6) [−vocalic] → [+vocalic] / {C,#} _ {C,#}
(glides vocalize when not adjacent to a vowel, e.g. /katw/ → [ka.tu])

Each of these rules stands in a potentially-bleeding relation to the other. To illustrate, consider an input /dahw/, which meets the structural description of both rules, as it has an /h/ not followed by a vowel, and a glide not adjacent to a vowel. Suppose that vocalization is ordered before deletion. Vocalization will convert /dahw/ into /dahu/, which no longer meets the structural description of /h/-deletion, as the /h/ is now pre-vocalic. Vocalization will have bled deletion. On the other hand, if /h/-deletion is ordered first, it will convert /dahw/ into /daw/, which no longer meets the structural description of vocalization, since the /w/ is now adjacent to a vowel. Deletion will have bled vocalization.

But now suppose that we permit the two rules to apply simultaneously. The input /dahw/ meets the structural description of both rules, so both will apply, giving [da.u]. In this scenario, the two rules are interacting in a *mutually counterbleeding* fashion. The glide is vocalizing even though /h/-deletion is taking away the C_ part of the environment for vocalization, and the /h/ is deleting even though vocalization is taking away the _[−voc] part of the environment for deletion.

Arguments involving mutual counterbleeding have also arisen in regards to debates between different approaches to opacity in Optimality Theory (Prince & Smolensky 2004 [1993]). Kiparsky (2001), for instance, demonstrates that Sympathy theory (McCarthy 1999) can produce mutual-counterbleeding interactions, which he uses as an argument in favor of Stratal OT (Kiparsky 2000, among many others) relative to Sympathy theory. Baković (2007b) points out that the two-level constraints proposed in McCarthy (1996) can, like simultaneous rule application, produce mutual counterbleeding interactions (or "mutually assured destruction" as Baković dubs this mode of process-interaction). Because OT approaches to opacity are not formally isomorphic to rule-based phonology with sequential, one-at-a-time application, we would not in general expect that the kinds of opacity that are modelable under any given OT approach to be identical to those which are modelable using sequential rule-application (Itô & Mester 2003). As such, for any given approach to opacity in OT, it is important that we enquire into its ability or inability to model 'exotic' and possibly unattested kinds of opaque interactions, such as mutual counterfeeding and mutual counterbleeding.

OT with Candidate Chains (OT-CC: McCarthy 2007) is a particularly interesting theory to examine in this regard. OT-CC is a theory in which candidates are (approximately) multi-step derivations, and so it bears a nearer resemblance to sequential, rule-based theories than many other approaches to opacity in OT. On the other hand, because OT-CC involves evaluating entire candidate derivations, the PREC constraints which regulate the order of processes in the derivation function as a type of global derivational constraint (or 'global rule': Lakoff 1970)—restrictions on the application of a process at one point in the derivation which can refer to conditions at non-adjacent (earlier or later) stages of the derivation. Allowing global rules brings with it the worry that the space of predicted opaque interactions would become completely unlimited, since "[a]ny imaginable rule can be described as a 'constraint on derivations'" (Chomsky 1972: 133-134). The types of global rules within OT-CC's known powers includes interactions whose existence is widely accepted, such as nonderived environment blocking (Wolf 2008: ch. 4), as well as others whose status is less clear, such as obligatory counterbleeding (Wolf 2008: §4.4.1) and counterfeeding from the past (Wilson 2006; cf. Wolf 2010). Given this, it becomes interesting (and important) to ask what limits, if any, exist on the kinds of derivational interactions that are predicted possible in OT-CC, and how those predictions match up to attested language typology.

This paper's goal is to provide a partial answer to that question, focusing on the two hypothetical interactions discussed above: mutual counterfeeding and mutual counterbleeding. As we'll see, the original version of OT-CC permits neither of these types of interactions; however, we will also see that some (but crucially not all) types of each could be brought into the fold via rather simple changes to the original OT-CC model. As the two changes are logically independent as to whether adopt them or not, we will end up considering four different versions of OT-CC, which differ in subtle ways about which kinds of opaque interactions they do and don't admit. Along the way, we will also look at the available evidence regarding the possible existence of mutual counterfeeding and of mutual counterbleeding. A few possible cases are reported, but probably none can yet be called unambiguously convincing.

This paper is organized as follows. Section 2 reviews the premises of OT-CC and the reasons for its ability to model global rules. Section 3 discusses mutual counterfeeding, along with the related interaction of self-counterfeeding. Section 4 discusses mutual counterbleeding, and section 5 summarizes the paper's overall conclusions.

2. Basic premises of OT-CC

In this section I will briefly review the fundamental assumptions of OT-CC which are essential to what follows.[4] In classic OT (Prince & Smolensky 2004 [1993]; though cf. their exploration of Harmonic Serialism in §5.2.3.3), each candidate is a direct mapping from an input to a candidate output form with no intermediate stages. OT-CC is different in that the candidate-generating function GEN produces candidates as gradual, one-step-at-a-time mappings from the input to a candidate output. These *candidate chains* are subject to three universal conditions that define what is a well-formed candidate:[5]

(7) *Gradualness*: Each form in the chain may differ from the previous one only by the performing of a single localized unfaithful mapping (LUM).

(8) *Harmonic improvement*: Each form in the chain must be more harmonic than the previous one, given the constraint hierarchy that prevails in the language in question.

(9) *Local Optimality*: Let $<\ldots f_{n-1}, f_n>$ be a valid chain in some language *L*. Let $g_1, \ldots g_m$ be all of the forms which can be formed from f_n by applying some LUM of the same type *T*. The chain $<\ldots f_{n-1}, f_n, g_i>$ is then a valid chain in *L* iff: (a) g_i is more harmonic than f_n; and (b) g_i is the most harmonic member of the set $\{g_1, \ldots g_m\}$.

(Informally: starting from any valid sub-chain, GEN can pursue only the single best way of making a change of type *T* to that subchain[6].)

With respect to the gradualness requirement (7), we can entertain any number of different hypotheses about what the inventory of available LUMs consists of. In this paper, I will follow McCarthy (2007a) in taking the available LUM types to be: delete one segment; epenthesize one segment; change one feature-value of one segment; metathesize two adjacent segments.

To briefly illustrate how the gradualness and harmonic improvement requirements interact, consider a language with the following rankings:

---

[4] A bibliography of works in or about OT-CC and Harmonic Serialism can be found at http://works.bepress.com/cgi/viewcontent.cgi?article=1101&context=john_j_mccarthy.

[5] This is setting aside the device of chain merger, which does not play a role in any of the issues discussed in this paper.

[6] A reviewer points out that, in principle, there is not necessarily a *unique* best: two or more forms produced via LUMs of the same type may tie in harmony. While formally possible, in practice ties are unlikely in OT, because the richness of the constraint-set CON makes it quite implausible that no constraint at all, however low-ranked, would differentiate between any two given candidates. That said, if there were a tie for Local Optimality between two forms g_i and g_j, we can imagine different hypotheses about how this would be dealt with. Perhaps the chain ending g_i and the chain ending g_j would both get to be included in the candidate set, or perhaps one or the other would be chosen randomly (as in the proposals by Hammond [1994] and Grimshaw [1997] that variation arises from multiple candidates tying for optimality).

(10) Final-C >> Max-V, *VoicedCoda >> Ident(voice) >> *VoicedObst

Final-C = One violation mark for every PWd that ends in a vowel
(McCarthy & Prince 1994)

Descriptively, this is a language with both coda devoicing and vowel apocope. For an input like /pada/, the following are the valid chains involving violation of Max-V and/or Ident(voice)[7]:

(11)
a. <pa.da> (do nothing)
b. <pa.da, pad> (delete final vowel)
c. <pa.da, pad, pat> (delete final vowel; devoice final consonant)

Let's compare these with a couple of examples of invalid chains. The hypothetical chain **<pa.da, pat> is invalid[8] under the gradualness requirement, given our assumption that deleting a segment and changing a segment's value of [voice] are distinct LUMs. This candidate flouts gradualness by performing two basic operations (two LUMs) in one go. Another invalid chain is **<pa.da, pa.ta, pat>, which runs afoul of the harmonic improvement requirement. The first step, of devoicing the onset /d/, is not harmonically improving with respect to the constraints included in the ranking in (10). The form [pa.da] violates *Voiced Obstruent, which [pa.ta] does not, but [pa.ta] gets rid of that violation at the expense of trading it in for a violation of the higher-ranked Ident(voice).

When a set of candidate chains like those in (11) compete, the markedness constraints assess only the final form in the chain. With only markedness and faithfulness constraints, we thus expect OT-CC to be just like classic, fully parallel OT to the extent that candidates with transparent interaction of processes would generally beat candidates corresponding to opaque interactions. For example, compare (11c) to (11b). Chain (11c), <pa.da, pad, pat> corresponds to a transparent interaction in which apocope feeds final devoicing. Chain (11b), <pa.da, pad> corresponds to an opaque interaction, in which apocope counterfeeds final devoicing. Given only the constraints in (10), the transparent chain (11c) will win, since the highest-ranked constraint among those depicted which has a preference between these two candidates is *VoicedCoda. It prefers (11c), which has a voiceless coda, over (11b), which has a voiced one.

[7] Here and throughout I'm assuming that resyllabification happens again for free at every step. The reason for assuming this in a Harmonic Serialism or OT-CC context is that deletion and epenthesis processes can be harmonically improving by virtue of producing less marked syllable structures, and can be blocked (i.e., can be harmonically disimproving) if they would produce marked syllable structure (see McCarthy 2010 for a fuller presentation of the argument). However, cf. Elfner (2008) for arguments that syllabification operations should be derivational steps in their own right.

[8] Following the conventions introduced in McCarthy (2007), I use a single asterisk to indicate a valid chain which is not the winning candidate, and a double asterisk to indicate a hypothetical chain which is invalid under one or more of the requirements (7-9).

To be able to get opaque candidates like (11b) to win, OT-CC augments the markedness and faithfulness constraints with an additional class of constraints which make demands about the order in which operations occur in the chains. These are known as PRECEDENCE constraints and have the following form:

(12) PREC(A,B)
Assign a violation-mark for every time that:
a. An operation which violates basic faithfulness constraint B occurs, and it is not preceded by an operation which violates basic faithfulness constraint A.

b. An operation which violates basic faithfulness constraint B occurs, and it is followed by an operation which violates basic faithfulness constraint A.

Again following McCarthy (2007), I will take the 'basic' faithfulness constraints to be the non-positional constraints of the MAX, DEP, IDENT, and LINEARITY families.

For our hypothetical language with apocope and final devoicing, if we wanted candidate (11b), with the counterfeeding interaction, to win, we could achieve this by ranking PREC(IDENT(voice), MAX-V) over *VOICEDCODA:

(13)

/pada/	FINAL-C	MAX-V	PREC (ID(voi), MAX-V)	*VOICOD	IDENT(voi)	*VOIOBST
a. <pa.da>	W1	L	L	L		1
b. ☞ <pa.da, pad>		1	1	1		1
c. <pa.da, pad, pat>		1	W2	L	W1	L

The winning candidate (13b) violates the PREC constraint once, because it has an instance of apocope (which violates MAX-V) which is not preceded by an instance of devoicing. This contravenes clause (a) of the PREC constraint. Its competitor (13c), however, gets two violation-marks from the PREC constraint, because it violates both clauses: its MAX-V-violating step fails to be preceded by an IDENT(voice)-violating LUM, and is followed by an IDENT(voice)-violating LUM.

A constraint PREC(A,B) is analogous (though not identical) to an extrinsic rule-ordering statement in rule-based phonology, in this case a statement that rule A (devoicing, in the hypothetical example just given) precedes rule B (apocope). The difference between PREC constraints and rule-ordering statements is that PREC constraints come into play 'after the fact': GEN initially constructs all candidate derivations compatible with the principles in (7-9), during which the PREC constraints have no role in harmonic evaluation. They only become active when the completed set of candidate derivations are compared.

The way in which OT-CC obtains opaque effects by constructing multiple alternative derivations and then comparing them places it in contrast with the related theory of Harmonic Serialism (Prince & Smolensky 2004 [1993]: §5.2.3.3). In HS, the output of one optimization is resubmitted to GEN to serve as the input to a new optimization with the same grammar. This looping through GEN and EVAL continues repeatedly until convergence: that is, until some input yields a fully-faithful output (which will occur once it is no longer harmonically improving to make any further changes). While in principle this looping could be coupled with an unrestricted GEN which could produce candidates that differed from the input in multiple ways, empirically HS is uninteresting unless coupled with the assumption that candidate outputs can differ from the input only in limited, basic ways (McCarthy 2000). HS is thus quite like OT-CC, except that only a single derivational path is constructed; the form reached when that path converges becomes the output. Because they both use a GEN which is limited to one 'basic operation' at a time, OT-CC and HS both exclude various kinds of unattesedly global interactions. This is because every process that occurs must be harmonically-improving at the point it occurs; nothing can occur which contributes to increased harmony only in conjunction with something else that happens later. There is now a substantial literature on how HS, coupled with particular assumptions about what the basic operations are, yields desirable results of this kind in various domains (see the bibliography referred to in footnote 4 for references.). These results will generally carry over into OT-CC as well, to the extent that similar assumptions are made about the operations available to GEN.[9]

One place where HS and OT-CC crucially differ is in the domain of opacity. In general, neither counterfeeding nor counterbleeding opacity are modelable in HS (McCarthy 2000).[10] For the counterfeeding scenario just considered, the problem is that when the form [pad] (produced via apoocope) is resubmitted to the grammar as an input, the winner will be [pat], with final devoicing:

(14)

/pad/	FINAL-C	MAX-V	*VOICOD	IDENT(voi)	*VOIOBST
a. ☞ [pat]				1	
b. [pad]			1		1

Because there is no way for the grammar to know that the [pad] presented to it ends in a voiced obstruent which became word-final via apocope (as opposed to being underlyingly final), there is no way to stop devoicing from applying here, short of ranking IDENT(voice) over *VOICOD, which will prevent coda devoicing from ever happening in this language. OT-CC does not have this limitation because in the final evaluation of derivational paths, the PREC constraints get to inspect entire derivations, so they can tell whether one process has applied before or after some other process.

[9] This means that, when opacity is not at stake, it is possible to investigate these locality issues using HS alone, without including the elaborations of OT-CC which make the task of analysis more complex.

[10] HS can, however, deal with at least some cases of opacity where the *way* a process applies is rendered opaque by some subsequent process—see Elfner (2008) on opaque stress/epenthesis interactions in HS, where the choice of location for stress is rendered opaque by subsequent vowel epenthesis.

In this way, OT-CC's approach to opacity echoes certain ideas about the interaction and ordering of rules which were debated in the 1970s. Postal (1972: 140-141) and Kisseberth (1973) observe that extrinsic rule-ordering statements can be regarded as a type of global derivational constraint (or 'global rule') in the sense of Lakoff (1970): they dictate that a certain operation cannot be performed at a certain point in a derivation if certain other operations either were or were not performed earlier in the derivation. OT-CC implements essentially this idea in order to get opacity to be possible, by setting up the PREC constraints and making each candidate's entire derivational history available to their assessment. Doing things in this way, though, means that OT-CC also opens the door to various types of 'global' interactions which are impossible (or at least, not straightforward) to model using the standard assumption of post-*SPE* rule-based phonology that rules apply sequentially, one at a time, and are Markovian (i.e., unable to "see" derivational history earlier than the form which is the output of the previous rule). This has been cited as both a virtue and as a liability of OT-CC. Wolf (2008: ch. 4), for instance, shows that OT-CC can be applied to the analysis of nonderived environment blocking (Kenstowicz & Kisseberth 1970; Kiparsky 1973) which is the classic example of a global rule.[11] In an NDEB scenario, a rule applies in derived environments but not in underived ones. NDEB is a global-rule effect because reference to prior derivational history is necessary to be able to tell derived environments apart from underived ones. On the other hand, Wilson (2006) shows that OT-CC can also produce a type of interaction which he dubs 'counterfeeding from the past', and which he argues is not attested.[12] Because the 'global' nature of PREC constraints allows OT-CC to potentially model forms of opacity which the standard rule-based approach cannot, establishing what limits (if any) exist on OT-CC's global powers is an important task.[13]

3. Mutual counterfeeding and self-counterfeeding

3.1 Mutual counterfeeding in OT-CC

Let us return, then, to the hypothetical mutual counterfeeding scenario described earlier and determine whether we can model it in OT-CC. The first thing we would need to make it possible is a ranking of markedness and faithfulness constraints which will

[11] Arguments for the existence in phonology of global rules of other types can be found in Lakoff (1972), Pyle (1972), McCawley (1973), Miller (1973, 1974), Kisseberth & Abasheikh (1975), and Underhill (1976), among others. The preceding list excludes proposed global rules which involve looking *ahead* to surface-structure conditions (e.g. Hill 1970; Kisseberth 1970), or which enforce base-reduplicant identity requirements (e.g. Wilbur 1973), which generally have a rather more mundane status in an OT context (see esp. McCarthy & Prince 1995 on the latter).

[12] However, for possible examples of counterfeeding from the past, see Wolf (2010) and the references therein. The Faroese data discussed by Anderson (1974: 167-174) also look like an example. See Odden (2008) for a demonstration that Sympathy (McCarthy 1999) also admits this form of opacity.

[13] If phonetics is a separate module fed by phonology, and the phonology is an OT-CC grammar, a second source of derivational-history effects may exist in addition to that opened up by the PREC constraints. Specifically, if the winning candidate (the output of the phonology) is a multi-step derivation, then phonetic interpretation may be able to refer to intermediate derivational steps. Gouskova & Hall (2009) propose an analysis of incomplete neutralization along these lines.

make the desired winning derivations <ehtəmu, etəmu> and <ahəpi, ahpi> harmonically improving. A straightforward way to do this with relatively general constraints would be the following:

(15) *Deletion of preconsonantal /h/ is harmonically improving*

/ehtəmu/	*hC	Max-h
etəmu *Is more harmonic than:*		*
ehtəmu	*	

*hC = One violation-mark for every sequence of [h] followed by a consonant
Max-h = One violation-mark for every input /h/ lacking an output correspondent.

(16) *Deletion of schwa is harmonically improving, even when it creates a [hC] cluster*

/ahəpi/	*ə	*hC	Max-ə
ahpi *Is more harmonic than:*		*	*
ahəpi	*		

*ə = One violation mark for every [ə].
Max- ə: One violation-mark for every input /ə/ lacking an output correspondent.

Additionally, the ranking *CCC >> *ə will ensure that schwa deletion will not be harmonically improving when it would produce a triconsonantal cluster:

(17)

/ehtəmu/	*CCC	*ə
ehtəmu *Is more harmonic than:*		*
ehtmu	*	

*CCC = One violation-mark for every sequence of three consecutive consonants.

This gives us the following overall rankings:

(18) *CCC >> *ə >> *hC >> Max-h
*ə >> Max-ə

Our next step is to identify, for the two inputs /ehtəmu/ and /ahəpi/, which other chains will be harmonically-improving with respect to the rankings just presented. These are the competitors which <ehtəmu, etəmu> and <ahəpi, ahpi> will have to beat in order to emerge as the winners. The chains we need to consider are as follows (assuming for expository simplicity that /h/-deletion and /ə/-deletion are the only phonological processes which occur in our hypothetical language):

(19) Chains for input /ehtəmu/:

<ehtəmu>	Do nothing
<ehtəmu, etəmu>	Delete /h/
<ehtəmu, etəmu, etmu>	Delete /h/, delete /ə/

(20) Chains for input /ahəpi/:

<ahəpi>	Do nothing
<ahəpi, ahpi>	Delete /ə/
<ahəpi, ahpi, api>	Delete /ə/, delete/h/

To get the mutual counterfeeding scenario, we need the second candidate listed for each input to beat the other two candidates for that input. The ranking conditions necessary for each winner to beat each of its losing competitors are expressed in the table of Elementary Ranking Conditions (ERCs: Prince 2002a,b) in (21) below. An ERC is a proposition about constraint ranking expressed as a winner/loser pair and a set of constraints, with each constraint annotated as to whether it favors the winner (indicated by a W), favors the loser (indicated by an L) or is indifferent. The proposition that this expresses is that for each row (each winner/loser pair), all of the L-assigning constraints have to be dominated by at least one of the W-assigning constraints (because otherwise the loser would win). The ERC table in (21) includes all of the markedness and faithfulness constraints already mentioned in this section, as well as the two Prec constraints which are potentially relevant, namely the two which mention only the two basic faithfulness constraints Max-h and Max-ə.

In order to depict the full set of ranking conditions necessary for the success of some analysis in OT-CC, we need to consider not only the ranking conditions which make a winning candidate beat each of its losing competitors; we need also to include the ranking conditions which will ensure that the processes of interest are or are not harmonically improving in the appropriate environments. In the case at hand, ensuring the harmonically-improving status of /ə/-deletion and /h/-deletion is equivalent to ensuring that the two chains <ahəpi, ahpi> and <ehtəmu, etəmu> each beat their fully faithful competitors.[14] The only 'extra' thing we need to include, then, is a row giving the ranking conditions to ensure that schwa deletion is not harmonically-improving if a CCC cluster would result. This is given in the last row of the ERC table, with [ehtəmu] as the 'winner' and [ehtmu] as its 'losing' competitor. I use '>' rather than '~' to separate the two forms being compared, to indicate the different status of this pair (as the

[14] This is not precisely true, since Prec constraints are at work in chain comparison but not in determining whether processes in chains are harmonically improving. To show only the ranking conditions needed for harmonic improvement, we would want to include ERC rows identical to those for the winner~loser pairs <ahəpi, ahpi> ~ <ahəpi> and <ehtəmu, etəmu> ~ <ehtəmu> which omit the Ls which indicate one of the Prec constraints' preference for the losers in these pairs. However, we need not actually include these, since the two rows with the Ls from the Prec constraints each entail an identical ERC row without the L, by the implicational rule of L-retraction (Prince 2002). That is, the rows without the Ls from the Prec constraints can be ignored because they contribute no additional information about what rankings are necessary, beyond what is conveyed by the rows with those Ls left in.

member which needs to be judged more harmonic, namely [ehtəmu], is not a winning surface form).[15]

(21)

W ~ L	*ə	*CCC	*hC	Max-ə	Max-h	Prec (Max-ə, Max-h)	Prec (Max-h, Max-ə)
<ahəpi, ahpi> ~ <ahəpi, ahpi, api>	(0~0)	(0~0)	L (1~0)	(1~1)	W (0~1)	(0~0)	W (1~2)
<ahəpi, ahpi> ~ <ahəpi>	W (0~1)	(0~0)	L (1~0)	L (1~0)	(0~0)	(0~0)	L (1~0)
<ehtəmu, etəmu> ~ <ehtəmu, etəmu, etmu>	L (1~0)	(0~0)	(0~0)	W (0~1)	(1~1)	W (1~2)	(0~0)
<ehtəmu, etəmu> ~ <ehtəmu>	(1~1)	(0~0)	W (0~1)	(0~0)	L (1~0)	L (1~0)	(0~0)
ehtəmu > ehtmu	L (1~0)	W (0~1)	(1~1)	W (0~1)	(0~0)	(0~0)	(0~0)

We can now use recursive constraint demotion (Tesar 1995; Tesar & Smolensky 2000; Prince 2002) to check whether these ERCs are consistent. We first look for constraints which assign no Ls, which are safe to install in the topmost ranking stratum. *CCC is the only one, and the W it supplies in the final row lets us eliminate that 'winner'~'loser' pair from consideration when we move on to the next pass of RCD:

(22)

W ~ L	*ə	*CCC	*hC	Max-ə	Max-h	Prec (Max-ə, Max-h)	Prec (Max-h, Max-ə)
<ahəpi, ahpi> ~ <ahəpi, ahpi, api>	(0~0)	(0~0)	L (1~0)	(1~1)	W (0~1)	(0~0)	W (1~2)
<ahəpi, ahpi> ~ <ahəpi>	W (0~1)	(0~0)	L (1~0)	L (1~0)	(0~0)	(0~0)	L (1~0)
<ehtəmu, etəmu> ~ <ehtəmu, etəmu, etmu>	L (1~0)	(0~0)	(0~0)	W (0~1)	(1~1)	W (1~2)	(0~0)
<ehtəmu, etəmu> ~ <ehtəmu>	(1~1)	(0~0)	W (0~1)	(0~0)	L (1~0)	L (1~0)	(0~0)
ehtəmu > ehtmu	L (1~0)	W (0~1)	(1~1)	W (0~1)	(0~0)	(0~0)	(0~0)

Constraints ranked so far: *CCC

On the next pass, RCD will crash, because all of the remaining constraints assign at least one L. Given this set of constraints, at least, our mutual counterfeeding scenario cannot be modeled in OT-CC.

[15] In each cell, the pair of integers *x~y* gives the numbers of violation-marks incurred by both candidates of the winner~loser pair in that row: the winner incurs *x* violations and the loser *y* violations.

What changes could we make to OT-CC in order to remove the inconsistency and make the mutual counterfeeding scenario modelable? Let us consider the structure of Prec constraints and the sources of the Ws and Ls that they assign in our table of ERCs. As originally formulated in McCarthy (2007a), a constraint Prec(A, B) has two clauses setting forth when violation-marks are to be assigned; the canonical Prec schema is here repeated from (12) earlier:

(23) Prec(A,B)

Assign a violation-mark for every time that:

a. An operation which violates basic faithfulness constraint B occurs, and it is not preceded by an operation which violates basic faithfulness constraint A.

b. An operation which violates basic faithfulness constraint B occurs, and it is followed by an operation which violates basic faithfulness constraint A.

Consider now the constraint Prec(Max-h, Max-ə). This constraint will assign a mark if (a) /ə/-deletion occurs and is not preceded by /h/-deletion, or (b) /ə/-deletion occurs and is followed by /h/-deletion. For the winner ~ loser pair <ahəpi, ahpi> ~ <ahəpi, ahpi, api>, both chains incur a violation of clause (a), the 'preceded by' clause, since each begins by deleting a schwa without having earlier deleted an /h/. However, the loser in this pair also violates clause (b), the 'not followed by' clause, since the loser deletes an /h/ after having previously deleted a schwa. It is by virtue of the loser's violation of clause (b) that Prec(Max-h, Max-ə) crucially assigns a W with regard to this winner ~ loser pair. Here, Prec(Max-h, Max-ə) is doing the work that needs doing in order to bring about mutual counterfeeding: it demands that we not do /h/-deletion if the conditions for /h/-deletion's application have been brought about by earlier application of schwa syncope.

The constraint Prec(Max-h, Max-ə) is not indifferent in one other case: the winner ~ loser pair <ahəpi, ahpi> ~ <ahəpi>. Here, the loser vacuously satisfies both clauses by virtue of having no Max-ə-violating LUM (and, indeed, no LUMs at all). However, the winner violates clause (a), the 'preceded by' clause, since it has a LUM of schwa-deletion which is not preceded by an earlier LUM of /h/-deletion. As in the previous pair we considered, the winner does not violate clause (b), the 'not followed by' clause, because in the winner the Max-ə-violating LUM of schwa syncope is not followed by /h/-deletion or by anything else. In this pair, then, it is the winner's violation of clause (a) which results in the L.

The situation is the same, *mutatis mutandis*, for Prec(Max-ə, Max-h) with respect to the chains derived from input /ehtəmu/: for one pair, the constraint supplies a W due to the loser's violation of the 'not followed by' clause, but for another pair, the constraint supplies an L due to the winner's violation of the 'preceded by' clause. By virtue of assigning an L which cannot be disposed of by ranking *CCC on the first pass of RCD, both Prec constraints are unrankable on the second pass, contributing to the inconsistency of the ERCs which we witnessed above.

The reader may by this point have guessed the strategy I will now explore for eliminating the inconsistency: split the Prec constraints in two, so that their (a) and (b) clauses are separate, independently-rankable constraints. In illustrating this option, I will call the spun-off (a)-clause A←B (read 'B implies preceding A'): it gives a violation mark if there is a B-violating LUM which is not preceded by an A-violating LUM. I will call the (b)-clause of Prec(A,B) *B-then-A: it gives a violation mark if an A-violating LUM follows a B-violating LUM.

If we split the Prec constraints in this way and then run RCD on the revised table of ERCs, RCD no longer crashes, and instead returns the ranking depicted in the following tableau. This shows that that mutual-counterbleeding ERCs, which were inconsistent under unsplit Prec, become consistent under the split-Prec alternative:

(24)

W ~ L	*CCC	*Max-h-then-Max-ə	*Max-ə-then-Max-h	*ə	*hC	Max-ə	Max-h	Max-ə ←Max-h	Max-h ←Max-ə
<ahəpi, ahpi> ~ <ahəpi, ahpi, api>	(0~0)	(0~0)	W (0~1)	(0~0)	L (1~0)	(1~1)	W (0~1)	(0~0)	(1~1)
<ahəpi, ahpi> ~ <ahəpi>	(0~0)	(0~0)	(0~0)	W (0~1)	L (1~0)	L (1~0)	(0~0)	(0~0)	L (1~0)
<ehtəmu, etəmu> ~ <ehtəmu, etəmu, etmu>	(0~0)	W (0~1)	(0~0)	L (1~0)	(0~0)	W (0~1)	(1~1)	(1~1)	(0~0)
<ehtəmu, etəmu> ~ <ehtəmu>	(0~0)	(0~0)	(0~0)	(1~1)	W (0~1)	(0~0)	L (1~0)	L (1~0)	(0~0)
ehtəmu > ehtmu	W (0~1)	(0~0)	(0~0)	L (1~0)	(1~1)	W (0~1)	(0~0)	(0~0)	(0~0)

Notice that this ranking is compatible with the rankings arrived at in (18) as being necessary to ensure harmonic-improvement in the desired winning chains, plus placing on top the *B-then-A constraints which produce the mutual-counterfeeding effect: neither process is permitted to apply if its application is crucially preceded by application of the other process.

We have thus demonstrated that, if the two clauses of the Prec constraints are split apart into independent constraints, the ERCs for our mutual counter-feeding scenario become consistent, meaning that OT-CC would now predict this to be possible pattern in natural languages. By contrast, if we keep the two clauses together as single constraints, the mutual-counterfeeding interaction remains predicted to be impossible. The reason for this impossibility can be described as follows. As we just said, in order to get either /ə/-deletion to be blocked in environments derived by /h/-deletion, we would need to rank PREC(MAX-ə, MAX-h) and its 'no MAX-ə violation after MAX-h violation' requirement above the markedness constraint *ə which motivates /ə/-syncope. However, this ranking also means that the other clause of PREC(MAX-ə, MAX-h)—the 'MAX-h violation must be preceded by MAX-ə violation' requirement—will be ranked above *ə and therefore by transitivity above *hC, the markedness constraint that motivates /h/-deletion. This means that /h/-deletion will be blocked in environments *not* derived by /ə/-syncope, which is precisely the place where we need /h/-deletion to be allowed.

3.2 Do mutual-counterfeeding interactions exist?

Excluding exchange rules (discussed below), I know of only one analysis which, if accepted, would involve countenancing mutual counterfeeding. This case—of which the hypothetical interaction in 3.1. is a simplification—involves schwa syncope and VN coalescence in Hindi-Urdu, as analyzed by Narang & Becker (1971). The facts are as follows. Schwa deletes except when it would create a triconsonantal cluster, accounting for alternations like:

(25)	/nikəl-naː/	[nikəlnaː]	'to come out'
	/nikəl-aː/	[niklaː]	'came out'

Narang & Becker (1971) observe that syncope also fails to occur in the environment ṼC_C:

(26) [ãːgən-õː], *[ãːgn-õː] 'courtyard-OBLIQUE.PL'

Based on this, they argue that nasal vowels in Hindi-Urdu are underlyingly /VN/ sequences, which coalesce to a nasal vowel before a consonant or word boundary.[16] If VN coalescence is ordered after syncope, then at the point syncope applies, the representation of 'courtyard-OBLIQUE.PL' is [angən-on], with the schwa in the environment CC_C and hence unable to syncopate. On this analysis, /VN/ coalescence counterfeeds syncope, because syncope creates strings like [ãːgən-õː] which meet the structural description of syncope (the schwa has no more than one consonant on each side), but which syncope nevertheless does not apply to.

Bhatia & Kenstowicz (1972) present a critique of Narang & Becker's (1971) analysis of the nasal vowels. Most pertinent to our concerns, they note that syncope also has the potential to feed /VN/ coalescence, specifically when deletion of a schwa would cause the nasal to become pre-consonantal, and hence eligible for coalescence. Ordering syncope before coalescence predicts that an underlying form like /maːnəsi/ 'mind-ADJECTIVAL' will first become /maːnsi/ via syncope, and then [mãːsi] via coalescence. This, however, is incorrect: the attested output is [maːnsi]. Indeed, Bhatia

[16] In the hypothetical mutual-counterfeeding interaction presented in the previous section, I used /h/-deletion in place of /VN/ coalescence in order to sidestep the complicating question of whether coalescence in OT-CC should be treated as one step or as separate steps of assimilation followed by deletion. See McCarthy (2007) for the latter view.

& Kenstowicz (1972) argue, coalescence never applies to /…VNC…/ sequences created by syncope from /…VNəC…/. If we accept Narang & Beker's (1971) analysis of nasal vowels in Hindi-Urdu as arising via coalescence from underlying /VN/, then coalescence and syncope will have to be mutually counterfeeding, which requires recourse to either global rules or simultaneous application.

Bhatia & Kenstowicz (1972) argue, however, that such theoretical innovations are probably not necessary in order to cope with the data at issue. Suppose we were to assume, contra Narang & Becker (1971), that vowel nasality in Hindi-Urdu is contrastive, that nasal vowels are simply nasal vowels rather than /VN/ sequences underlyingly, and that there is no synchronic rule of /VN/ coalescence. The failure of syncope to feed /VN/ coalescence is hardly surprising if the latter rule does not exist. The underlying status of vowel nasality, and the synchronic non-existence of the coalescence rule, are supported by several observations made by Bhatia & Kenstowicz (1972). First, there are a number of lexical exceptions to coalescence. They cite [ʃaːnt] 'quiet', [pɣaːnt] 'district', [kaːntaː] 'wife', [aːntɣik] 'internal', [xzaːnciː] 'treasurer', and [jaːnkiː] 'wife of lord Ram' as examples. Second, coalescence does not apply to /VN+C/ sequences created by morpheme concatenation:

(27)	[jaːn-kaːr], *[jãːkaːr]	'clever'
	[b^haːn-jaː], *[b^hãː-jaː]	'sister's son'

Third, they note that "many speakers of Hindi now syncopate [schwa] even if the preceding syllable contains a nasal vowel [i.e., in the environment ṼC_CV] […] In other words, nasalized vowels are becoming (and have become for many speakers) just like other vowels with respect to syncopation and hence one of the prime motivations for deriving them from [VNC] clusters has been lost, which further suggests that nasal vowels are phonemic" (p. 210). If these arguments go through, and there is no synchronic VN coalescence rule, then there is obviously no rule interaction (mutually-counterfeeding or otherwise) that needs to be accounted for here.

Besides the possible Hindi-Urdu example, there is one class of mutual-counterfeeding interactions which has been argued to exist, but which would remain unmodelable in OT-CC even if the Prec constraints were split apart. This is the mutual-counterfeeding interaction which would hold between the two halves of an exchange rule. Consider, for illustration, the following example:

(28) [α voice]→[−α voice] / _#

This rule will reverse the [±voice] specification of a word-final segment: /pat/ will become [pad] and /pad/ will become [pat]. This exchange rule is equivalent to the following two 'ordinary' rules interacting in a mutually-counterfeeding manner:

(29) [+voice]→[−voice] / _#
[−voice]→[+voice] / _#

If the two separate rules of final devoicing and final voicing apply one at a time, then whichever applies first will feed the other, resulting in the first rule's application being totally obscured (since the second rule will always undo its effects):

(30)

	/pad/	/pat/
Final devoicing	pat	pat
Final voicing	pad	pad

	/pad/	/pat/
Final voicing	pad	pad
Final devoicing	pat	pat

If, however, the two rules can be specified to apply simultaneously, and neither is permitted to apply again subsequently, then neither will apply to the output of the other, and we will get the flip-flop scenario where /pad/ goes to [pat] and /pat/ goes to [pad] (Chafe 1968: 124-125; Anderson 1974: 91-97). Even if we don't allow literally simultaneous rule application, then the alpha-variable notation allows us to sneak it in through the back door by packaging final voicing and final devoicing in a single rule, whose application then becomes equivalent to simultaneously attempting to apply both sub-rules.

This particular type of mutual-counterfeeding scenario cannot be recapitulated in OT-CC because the chains <pad, pat> and <pat, pad> cannot both be harmonically-improving given a single ranking of markedness and faithfulness constraints (Moreton 1999; see Kavitskaya & Staroverov [2010] for discussion in an OT-CC context). In an exchange-rule scenario, we would need a process motivated by markedness constraint *A* which created new violations of markedness constraint *B*, and a process motivated by markedness constraint *B* which created new violations of markedness constraint *A*; the problem is that the two processes are inconsistent in their requirements about the relative ranking of *A* and *B*. The hypothetical example we considered above, however, does not have this character: schwa syncope does indeed create new violations of *hC, thus giving schwa syncope the potential to feed /h/-deletion. However, /h/-deletion has the potential to feed schwa syncope not by creating new violations of *ə (the markedness constraint which motivates syncope) but instead by reducing the number of consonants to one side of a schwa, thus removing the barrier to syncope posed by high-ranked *CCC.

Given all this, we can conclude that even a version of OT-CC where the Prec constraints are split in half is less powerful than a version of rule-based phonology in which all rules apply simultaneously. The latter theory allows the exchange-rule type of mutual counterfeeding, but the former does not. The exclusion of exchange rules in OT-CC and in OT generally is typologically desirable, since very few if any plausible examples of input-output exchange processes have ever been reported. Most if not all reported examples of exchange rules can be regarded as morpholexical in nature, i.e. driven by the need to give overt realization to a morphological category (see e.g.

Anderson & Browne 1973; Moreton 1999). Whether even this type of exchange-rule exists has recently come in for much questioning, in particular with regard to the well-known example of final voicing alternations in DhoLuo (see e.g. Pulleyblank 2006; Trommer 2007). The other most famous example of an exchange process is the Taiwanese tone-sandhi circle; a number of arguments have been raised that this is no longer part of the productive phonology of the language (see esp. Myers & Tsay 2002; Zhang, Lai & Turnbull-Sailor 2006; and references therein).

3.3 Self-counterfeeding

As mentioned, a theory with all-simultaneous rule application does not permit any rule to feed any other. One subcase of this prediction is that no rule will ever be able to feed itself. One-at-a-time rule application makes the same prediction, unless provision is made for some or all rules to be iterative (Howard 1972; Johnson 1972; Anderson 1974, among others). In this section, I will briefly demonstrate that certain self-counterfeeding interactions can obtain in OT-CC under the same conditions as are required for mutual counterfeeding, namely the splitting in half of the PREC constraints. Unlike mutual counterfeeding, several plausible examples of self-counterfeeding are known. I will illustrate OT-CC's predictions using a schematic example of self-counterfeeding apocope where /...CVV#/ surfaces as [...CV#] and /...CV#/ surfaces as [...C#]. Chemehuevi (Press 1979: §1.33; Vago & Batistella 1982) is one language where this has been argued to occur:

(31) Non-iterating final vowel deletion in Chemehuevi

/moa/	[mo]	'father'
/pacɨ/	[pac]	'daughter'
/nukwivaa/	[nukwiva]	'will run'

Other languages which have been argued to have self-counterfeeding word-final shortening or deletion processes include Catalan (Wheeler 1979; Boersma 2001), Hidatsa (Harris 1942; Kenstowicz & Kisseberth 1977: 178-178, 1979: 318-319; McCarthy 2003), Karok (Bright 1957: §§321, 331), Latvian (Halle & Zeps 1966), Lithuanian (Lightner 1972), Odawa (Piggott 1975), Ponapean (Howard 1972: 179-81), and Woleaian (Sohn 1975).

For the two inputs /CVV/ and /CV/ that participate in an interaction of this sort, the chains of interest are:

(32) /CVV/:
<CVV>
<CVV, CV>
<CVV, CV, C>

(33) /CV/:
<CV>
<CV, C>

The harmonic-improvement requirement of OT-CC has an important consequence for how we analyze this scenario. The /CV#/→[C#] mapping can be straightforwardly attributed to FINAL-C. However, the /CVV#/→[CV#] mapping cannot be attributed to the same constraint. The forms [CVV#] and [CV#] both end in a vowel, so the unfaithful mapping brings no payoff in improved performance on FINAL-C. Therefore, there has to be some other markedness constraint which favors [CV#] over [CVV#]; a reasonable possibility is that [CV#] is preferred because it lacks the vowel hiatus found in [CVV#]. Now, in principle, we could attribute the harmonically-improving status of the /CV#/→[C#] and /CVV#/→[CV#] mappings to a single markedness constraint if that constraint assigned two or more marks to [CVV#], and some smaller but non-zero number of marks to [CV#]. For instance, if there were a constraint that said something like 'assign one violation-mark to every vowel that is not followed by a consonant in the same prosodic word', that constraint would assign two violation-marks to [CVV#], one to [CV#], and none to [C#].[17] Absent such a constraint, however, this means that the two steps of a 'mutual counterfeeding' chainshift will have to be attributed to different markedness constraints. This contrasts interestingly with how the scenario might be analyzed using rules, where a single, non-iterative rule that deleted a word-final vowel could be responsible for both the /CV#/→[C#] and /CVV#/→[CV#] mappings. The potential for 'mutual counterfeeding' to be modelable in OT-CC thus depends in delicate ways on the substantive contents of the constraint set: if we don't have a single markedness constraint available which would motivate both steps, then a second constraint (like *HIATUS in this example) would have to be found which could kick-start the chain shift.

Under the split-PREC hypothesis, supposing FINAL-C and *HIATUS to be the two pertinent markedness constraints, and MAX-V to be the pertinent faithfulness constraint, the following tableau shows the ERCs of the required winner/loser pairs, along with the ranking of the constraints found by RCD:

(34)

	*HIATUS	*MAX-V-THEN-MAX-V	FINALC	MAX-V	MAX-V←MAX-V
<CVV, CV> ~ <CVV>	W (0~1)	(0~0)	(1~1)	L (1~0)	L (1~0)
<CVV, CV> ~ <CVV, CV, C>	(0~0)	W (0~1)	L (1~0)	W (1~2)	(1~1)
<CV, C> ~ <CV>	(0~0)	(0~0)	W (0~1)	L (1~0)	L (1~0)

[17] Thanks to an anonymous reviewer for raising this point. I'm unaware of any proposals for the alternative formulation of FINAL-C given in the text, though it does bear a certain resemblance to Smith's (to appear) proposal that the constraint ONSET should be formulated to penalize syllables where the head segment of the syllable is not preceded by some other segment in the syllable.

Since a consistent ranking for these winner/loser pairs can be found, we conclude that self-counterfeeding is possible in OT-CC under the split-Prec hypothesis. We may now compare that result to what happens with unsplit Prec. Here the ERCs are as follows; to save space we proceed directly to the result of the first pass of RCD, where the constraint *Hiatus is installed in the topmost stratum:

(35)

	*Hiatus	FinalC	Max-V	Prec(Max-V, Max-V)
<CVV, CV> ~ <CVV>	W (0~1)	(1~1)	L (1~0)	L (1~0)
<CVV, CV> ~ <CVV, CV, C>	(0~0)	L (1~0)	W (1~2)	W (1~2)
<CV, C> ~ <CV>	(0~0)	W (0~1)	L (1~0)	L (1~0)

Constraints ranked so far: *Hiatus

After this, however, RCD will crash, since there is no remaining constraint that doesn't assign an L.

To conclude this section, then, separating the two clauses of Prec constraints into independent constraints results in both mutual-counterfeeding and self-counterfeeding being allowed. In both cases, though, there are limitations. For mutual-counterfeeding, the limitation is formal: the exchange-rule type of mutual counterfeeding remains unmodelable. For self-counterfeeding, the limitation is substantive: we need either a single markedness constraint which will motivate both steps of the chainshift, or separate markedness constraints to motivate each of the steps independently. These observations foreshadow considerations that we will encounter in the next section when considering the conditions under which OT-CC can model mutual counterbleeding interactions.

4. Mutual counterbleeding

4.1 Mutually-counterbleeding processes that violate different basic faithfulness constraints

To illustrate what is required to be able to analyze mutual counterbleeding in OT-CC, let us return to the hypothetical scenario presented in the introduction. Suppose that we have a language where /h/ deletes in coda position:

(36) /pah/ → [pa]
cf. /kuhi/ → [ku.hi], *[ku.i]

Suppose that in this same language, glides vocalize when not adjacent to a vowel :

(37) /atj/ → [a.ti]
cf. /kej/ → [kej]

Mutual counterbleeding interaction of these two processes would mean that word-final (or pre-consonantal) [...Vhj] sequences undergo both /h/ deletion and glide vocalization:

(38) /ahj/ → [a.i], *[aj], *[a.hi]

The two processes are mutually-counterbleeding in the /ahj/ → [a.i] mapping because deleting the /h/ places the glide next to a vowel (taking away the environment for vocalization), and vocalizing the glide would allow /h/ to syllabify as the onset to the newly-created vowel, taking away the environment for /h/-deletion.

For the inputs where only one of the two processes is applicable, defining what the relevant chains would be is straightforward:

(39) /pah/ <pa>
<pah, pa> (winner)

(40) /atj/ <atj>
<atj, a.ti> (winner)

Things become more complicated when we consider an input like /ahj/, which we want to undergo both processes. Obviously, the candidate set for this input will have to include the fully-faithful chain, as well as the chains where only one or the other of the two processes occurs:

(41) /ahj/ <ahj>
<ahj, aj>
<ahj, a.hi>

The complication is how we can get a chain that ends in [a.i], which is the intended winner for this input given our assumption that the language has mutual counterbleeding. To be able to get to [a.i], it must be the case that either (a) it is harmonically improving to change /aj/ to [a.i] via glide vocalization, or (b) it is harmonically improving to change /a.hi/ to [a.i] via /h/-deletion. To achieve the surface form corresponding to mutual counterbleeding interaction, it is necessary that one process happen first and the other second, and whichever goes second will thus have to be harmonically improving in more situations than our original description of the hypothetical language would apparently call for. For illustration, let us suppose that it is the change of /aj/ to [a.i] which is harmonically improving—that is, it is in fact always harmonically improving to change a glide to a vowel, presumably because a general markedness constraint *GLIDE outranks IDENT(vocalic). (The choice is arbitrary; the same issues of analysis laid out below would arise if we assumed instead that the change of /a.hi/ to [a.i] was harmonically improving.)

Allowing that intermediate /aj/ is less harmonic than unfaithful [a.i], however, obviously imperils our characterization of the language as vocalizing those glides which are not vowel-adjacent; the ranking of *GLIDE over IDENT(vocalic) means that it will also be harmonically improving to vocalize a glide which is underlyingly vowel-adjacent:

(42)	/aj/	<aj>	(intended winner)
		<aj, a.i>	

What we need to happen, then, is for <aj> to beat <aj, a.i> while also having <ahj, aj, a.i> beat <ahj, aj>. *GLIDE's preference for glide vocalization in all circumstances whatsoever will have to be over-ruled when the glide in question is *underlyingly* vowel-adjacent (as in the <aj> ~ <aj, a.i> comparison), but not when the glide in question is derivedly vowel-adjacent as a result of /h/-deletion (as in the <ahj, aj, a.i> ~ <ahj, aj> comparison). That is, the permissibility of glide-deletion has to be sensitive to derivational history, and so the constraint which crucially outranks *GLIDE therefore has to be a PREC constraint.

Now we are ready to proceed to the full analysis. Suppose that glide vocalization in non-vowel-adjacent contexts is due to a constraint against complex syllable margins (*COMPLEX). If we put *COMPLEX, *GLIDE, *CODA/h, the violated faithfulness constraints MAX-h and IDENT(vocalic), and the two PREC constraints relating those two LUMs into the table of ERCs for the relevant winner/loser pairs, and run RCD on those ERCs, we succeed in finding the following ranking:

(43)

	*COMPLEX	*CODA/h	PREC (MAX-h, ID(voc))	*GLIDE	MAX-h	IDENT (voc)	PREC (ID(voc), MAX-h)
<pah, pa> ~ <pah>	(0~0)	W (0~1)	(0~0)	(0~0)	L (1~0)	(0~0)	L (1~0)
<atj, a.ti> ~ <atj>	W (0~1)	(0~0)	L (1~0)	W (0~1)	(0~0)	L (1~0)	(0~0)
<aj> ~ <aj, a.i>	(0~0)	(0~0)	W (0~1)	L (1~0)	(0~0)	W (0~1)	(0~0)
<ahj, aj, a.i> ~ <ahj, aj>	(0~0)	(0~0)	(0~0)	W (0~1)	(1~1)	L (1~0)	L (2~1)
<ahj, aj, a.i> ~ <ahj, a.hi>	(0~0)	(0~0)	W (0~1)	(0~0)	L (1~0)	(1~1)	L (2~0)
<ahj, aj, a.i> ~ <ahj>	W (0~1)	W (0~1)	(0~0)	W (0~1)	L (1~0)	L (1~0)	L (2~0)

Since RCD succeeded in finding a ranking, this analysis obviously works; less obvious is to unpack exactly how and why it works. For an input like /aj/, with an underlyingly vowel-adjacent glide, the ranking *GLIDE over IDENT(voc) has the effect of favoring the vocalizing chain <aj, a.i> over faithful <a.i>. However, it is the faithful candidate that wins because *GLIDE is in turn dominated by PREC(MAX-h, ID(voc)). The vocalizing chain <aj, a.i> runs afoul of PREC(MAX-h, ID(voc)) because it vocalizes the glide without previously having deleted an /h/.

While vocalization of glides is thus blocked for inputs like /aj/, it is allowed for inputs like /ahj/, where the winning chain <ahj, aj, a.i> deletes its /h/ before performing vocalization. The constraint PREC(MAX-h, ID(voc)) thus has no objection to this chain, which is then able to beat its non-vocalizing competitor <ahj, aj> by virtue of the preference exerted by *GLIDE. Under the assumptions we are making, there is an inherent ordering relation between /h/-deletion and /j/-vocalization for an input like /ahj/ where the processes can interact. It is possible for glide vocalization to occur after /h/ deletion, given our assumption that glide vocalization is harmonically improving across the board. However, it is not possible to vocalize first and then delete the /h/, given our assumption that /h/-deletion is not harmonically improving except when the /h/ is a coda. Vocalization would bleed /h/-deletion, so if both are to happen, the order has to be /h/-deletion first, which subsequent glide vocalization then counterbleeds. The effect of the ranking PREC(MAX-h, ID(voc)) >> *GLIDE >> IDENT(voc) thus is that vocalization of glides is harmonically improving at any point during chain construction, but that glide vocalization is blocked in chain comparison when there isn't a preceding instance of /h/-deletion that the vocalization counterbleeds.

That isn't the final word, though, because PREC(MAX-h, ID(voc)) is itself outranked by *COMPLEX. In comparisons like <atj, a.ti> ~ <atj>, the winner has glide vocalization, even though there is no /h/-deletion for it to counterbleed. Violation of the PREC constraint is tolerated in cases like this where vocalization is needed in order to eliminate complex syllable margins on the surface. The overall effect of the ranking found is that glide vocalization can be done at any time during chain construction, but only happens in winning candidates when it either (a) serves to eliminate violations of *COMPLEX, or (b) counterbleeds prior application of /h/-deletion.

The analysis of an apparent mutual counterbleeding interaction in OT-CC thus demands that the interaction in fact be treated as a partially concealed case of *obligatory counterbleeding* (Wolf 2008: §4.4.1), that is, of a process which is allowed to happen only when it will counterbleed some other process. Except for the possible case of the interaction of spirantization and vowel shortening in Chimwi:ni (Kisseberth & Abasheikh 1975; Kenstowicz & Kisseberth 1977; Hyman 1993), examples of obligatory counterbleeding are not known to me. Indeed, it is precisely to rule out this form of opacity that McCarthy (2007) proposes the following ranking metaconstraint for OT-CC:

(44) B >> PREC(A,B)

Looking back at the ranking arrived at for the mutual counterbleeding scenario, we may observe that this ranking does not comply with the metaconstraint: *GLIDE has to outrank IDENT(vocalic) to get vocalization to be harmonically improving across the board, and PREC(MAX-h, ID(voc)) has to outrank *GLIDE in order to block vocalization when it would not counterbleed (and thus follow) /h/-deletion. Thus, by transitivity PREC(MAX-h, ID(voc)) has to outrank IDENT(vocalic)—a case of precisely what the ranking metaconstraint forbids.

While a version of OT-CC without the ranking metaconstraint will allow mutual counterbleeding by virtue of allowing obligatory counterbleeding, its abilities in that regard are not unlimited. As with self-counterfeeding, the limitations are substantive in nature. The limitation is imposed by the need for a markedness constraint which will play the role played by *GLIDE in the analysis above: a constraint which will favor performing one of the changes that make up the mutual-counterbleeding interaction across the board. If we can't find such a constraint, then an analysis of the kind we constructed above will not be possible. The following hypothetical case where a constraint of the requisite kind may not be available is inspired by some of the Hungarian facts discussed by Vago (1977: §2) in arguing in favor of the need for extrinsic rule ordering. Suppose that some language has a rule of regressive velar assimilation:

(45) C → [velar] / _[velar]
(e.g. /anka/ → [aŋka])

And suppose that the language also has a rule of progressive assimilation to retroflexivity (Steriade 2001):

(46) C → [retroflex] / [retroflex]_
(e.g. /aɳta/ → [aɳʈa])

What will happen, then, to an input like /eɖke/ which has a retroflex followed by a velar? Under one-at-a-time application, we expect either [eɖʈe] (if retroflex assimilation applies first) or [egke] (if velar assimilation happens first). With either order, the rule that goes second will be bled by the application of the first, since the source of assimilation for the second rule is destroyed by assimilating it to the place of articulation of the source of assimilation for the first rule.

But what if the two rules apply simultaneously? The input /eɖke/ meets the structural description of both rules, so we expect both to apply, each counterbleeding the other, to yield [egʈe]. Such an input-output mapping is exceedingly implausible. Indeed, in light of Steriade's (2001) perceptually-based account of why retroflex assimilation is progressive while assimilation in other C-place features is regressive—namely that retroflex place is best cued in postvocalic position while other places of consonant articulation are best cued in prevocalic position—the /eɖke/→[egʈe] mapping is strikingly perverse, since it means that both of the two C-place features are being shifted to a position where they are perceptually less well cued.

In terms of OT-CC, this hypothetical example of mutual counterbleeding is probably excluded because, once we perform one of the feature changes, it's unlikely that any markedness constraint is available which would motivate the other change.

Suppose that retroflex assimilation happens first, giving the subchain <eɖke, eɖʈe>. From here, it's hard to see what is to be gained in markedness terms by changing the first retroflex stop to a velar. (Place dissimilation might be a motivation, but then we would probably expect the first /ɖ/ to change into something less marked than [g].) Similarly, if velar assimilation happened first, giving the subchain <eɖke, egke>, it's difficult to see why we would want to change the /k/ into [ʈ]. The concoction of other, yet more absurd, mutual-counterbleeding scenarios of this sort may be left to the reader's imagination.

An interesting point to ask about the ranking metaconstraint at this juncture is: could the metaconstraint be retained under a version of OT-CC which split apart the Prec constraints, as discussed in section 3? The obligatory-counterbleeding type of interactions which the B >> Prec(A,B) ranking metaconstraint is meant to rule out arises from the activity of the 'B-violation must be preceded by A-violation' clause of Prec(A,B). If examples of mutual counterfeeding were indeed found, motivating a division of the clauses of Prec constraints, it would be possible for us to retain the ranking metaconstraint in modified form as B >> A←B. Recall that the A←B constraints ended up bottom-ranked in the mutual-counterfeeding analysis; the mutual-counterfeeding effect results from the high-ranked status of the *B-then-A constraints. Thus in OT-CC, the choice of whether we allow mutual counterfeeding (by splitting Prec) and of whether we allow mutual counterbleeding (by dropping the ranking metaconstraint) are logically independent. This is an interesting difference from assuming rule-based phonology and that all rules apply simultaneously; those assumptions will predict both mutual counterfeeding and mutual counterbleeding.

One last point before concluding this subsection: the considerations which forced us to assume an across-the-board process of glide vocalization which is subject to obligatory counterbleeding nicely illustrate why Harmonic Serialism by itself doesn't help with counterbleeding opacity (which was first shown by McCarthy 2000). For an input like /ahj/, either vocalizing the glide or deleting the /h/ would eliminate the violations of *h/coda and *Complex at a stroke; whether [a.hi] or [aj] won on the first pass through Gen and Eval would be decided by lower-ranked constraints. Let us suppose, arbitrarily, that [aj] is the winner due to Max-h being lower-ranked than Ident(vocalic):

(47)

/ahj/	*h/coda	*Complex	Ident(voc)	Max-h
a. [ahj]	W1	W1		L
b. [a.hi]			W1	L
c. ☞ [aj]				1

At this point, the HS treatment encounters the same problem we ran into in chain construction for the OT-CC analysis: given only the constraints depicted, there is no reason for [aj] to map to [a.i] on the next pass. The assumed markedness motivation for glide vocalization, namely violation of *Complex, was already taken care of by /h/-deletion. To be able to get [aj] to map to [a.i], we could assume, as before, that *Glide outranked IDENT(vocalic), but as before that will mean that underlying as well as derived /aj/ would undergo glide vocalization. In the OT-CC analysis, we foreclosed that via the ranking of the PREC constraint which imposed obligatory counterbleeding—a move which is not available in HS because there is only one derivational path, no PREC constraints, and no other mechanism for referring to derivational history.

4.2 Mutually-counterbleeding processes that violate the same basic faithfulness constraint

The hypothetical mutual counterbleeding scenario in the previous section involved two processes—/h/-deletion and glide vocalization—which violate different basic faithfulness constraints. What about mutual-counterbleeding scenarios in which the two processes violate the same basic faithfulness constraint? It turns out that modeling this sort of interaction in OT-CC only becomes possible if we both lift the ranking metaconstraint *and* split the PREC constraints in two. I'll illustrate using an example of a hypothetical mutual-counterbleeding interaction between rules of the following form:

(48) X→Ø /_X
X→Ø / X_

If these rules apply simultaneously, then two adjacent underlying Xes will *both* get deleted, while a solitary underlying X will survive intact. I will illustrate what is required to analyze an interaction of this sort in OT-CC using a hypothetical example in which the Xes are High tones. This example is a (severely) simplified version of an interaction which occurs in the Bantu language Tachoni (Odden 2008: §3.2.3; see the discussion in the next subsection for more details).

Suppose that in some language, underlying High tones which are not next to any other High surface intact, while underlying High-High sequences suffer the deletion of both High tones. Here are the chains that need to be dealt with:

For input /$H_{root}H_{affix}$/:

(49)	<HH>	(do nothing)	
	<HH, H0>	<MAX(H)@2>	
	<HH, H0, 00>	<MAX(H)@2, MAX(H)@1>	(winner)

For input /H/:

(50)	<H>	(do nothing)	(winner)
	<H, 0>	<MAX(H)@1>	

In the above, I'm using H to indicate a tone bearing unit that carries an High tone and 0 to indicate a toneless TBU. Additionally, because we have two potential loci for High-tone-deletion, we require in our sequences of LUMs some notation which disambiguates which is the one in question. For this purpose I follow the use of '@' in

McCarthy (2007): here, 'Max(H)@1' means deletion of the first High, and 'Max(H)@2' means deletion of the second High.

For input /HH/, we are intending for the winner to be [00], with both of the High tones being deleted. Starting from /HH/, in principle both /H0/ and /0H/ could be harmonically-improving steps that could follow in the chain. However, because of OT-CC's Local Optimality assumption, only one of these options can actually appear in the candidate set. Both /H0/ and /0H/ would be produced from /HH/ via performing the same basic operation of High-tone-deletion. Therefore, only the more harmonic of these two possible moves may actually be pursued in the process of constructing chains. I will assume in the case of (49) that it's /H0/ which is locally optimal relative to /0H/, due to the existence of root-faithfulness (McCarthy & Prince 1995; Casali 1997; Beckman 1998). Local Optimality has an important consequence for this miniature example: in the intended winner <HH, H0, 00>, the two High tone deletions will inevitably have to be ordered, since Local Optimality forces us to consider only one High tone deletion as the one that happens first, and so each High tone deletion cannot but be ordered with respect to the other.

Let us first consider what happens under the split-Prec hypothesis. The following table of ERCs includes the markedness constraints OCP(H) (motivating deletion in cases where there are two High tones), *H (motivating deletion of any High tone), the countervailing faithfulness constraints Max(H) and Max(H)$_{Root}$, and the two split-apart halves of Prec(Max(H), Max(H)). Given the ERCs with these constraints and the requisite winner/loser pairs, RCD succeeds in finding the following ranking:

(51)

	OCP(H)	Max(H) ←Max(H)	*H	Max(H)$_{Root}$	Max(H)	*Max(H)-then-Max(H)
<HH, H0, 00> ~ <HH>	W (0~1)	L (1~0)	W (0~2)	L (1~0)	L (2~0)	L (1~0)
<HH, H0, 00> ~ <HH, H0>	(0~0)	(1~1)	W (0~1)	L (1~0)	L (2~1)	L (1~0)
<H> ~ <H, 0>	(0~0)	W (0~1)	L (1~0)	W (0~1)	W (0~1)	(0~0)
H0 > HH	W (0~1)	(0~0)	W (1~2)	(0~0)	L (1~0)	(0~0)
0 > H	(0~0)	(0~0)	W (0~1)	L (1~0)	L (1~0)	(0~0)
H0 > 0H	(0~0)	(0~0)	(1~1)	W (0~1)	(1~1)	(0~0)

A ranking was successfully found; however, note that this requires that in addition to adopting the split-PREC hypothesis, we have to set aside the ranking metaconstraint, since we have MAX(H)←MAX(H) ranked above MAX(H).

Now suppose we attempt the exact same thing but with unsplit PREC. In that case we start with the following ERCs; to save space we proceed directly to showing the result of the first pass of RCD:

(52)

	OCP(H)	*H	MAX(H)	MAX(H)$_{Root}$	PREC (MAX(H), MAX(H))
<HH, H0, 00> ~ <HH>	W (0~1)	W (0~2)	L (2~0)	L (1~0)	L (2~0)
<HH, H0, 00> ~ <HH, H0>	(0~0)	W (0~1)	L (2~1)	L (1~0)	L (2~1)
<H> ~ <H, 0>	(0~0)	L (1~0)	W (0~1)	W (0~1)	W (0~1)
H0 > HH	W (0~1)	W (1~2)	L (1~0)	(0~0)	(0~0)
0 > H	(0~0)	W (0~1)	L (1~0)	L (1~0)	(0~0)
H0 > 0H	(0~0)	(1~1)	(1~1)	W (0~1)	(0~0)

Constraints ranked so far: OCP(H)

On the next pass, RCD will crash, as all the remaining constraints have at least one L.

So why exactly do we need split PREC constraints, in addition to freedom from the ranking metaconstraint, in order to cope with mutual counterbleeding interactions where the two processes violate the same basic faithfulness constraint? Let's examine where we got to after the initial pass of RCD under the unsplit-PREC hypothesis, as depicted in (52). PREC(MAX(H), MAX(H)), like all of the constraints, assigns both a W and an L, which is why RCD crashed. It assigns a W in the comparison <H> ~ <H, 0>, because the loser violates the MAX(H)←MAX(H) clause. (Inevitably, every candidate with one or more MAX(H) violations will do so at least once, because one of them will of necessity be the first one, and as it is preceded by no other MAX(H) violation, the MAX(H)←MAX(H) clause is violated.)

The L assigned by PREC(MAX(H), MAX(H)) which is left after OCP(H) has been ranked is with respect to the comparison <HH, H0, 00> ~ <HH, H0>. Both winner and loser here equally violate the MAX(H)←MAX(H) clause, for the reasons mentioned in the previous paragraph. However, the winner violates the "*MAX(H)-THEN-MAX(H)" clause by virtue of having two Max(H) violations which are ordered with respect to one another.

To get the analysis to work, we need the W resulting from the "MAX(H)←MAX(H)" clause's preference for <H> over <H, 0> to outrank *H's opposite preference between these two candidates. That is, we need this clause to be high-ranked in order to inhibit deletion of solitary High tones, except when deleting some solitary High tone will counterbleed earlier, OCP-driven deletion of another High. Contrariwise, we also need for the L resulting from the "*MAX(H)-THEN-MAX(H)" clause's preference for <HH, H0> over <HH, H0, 00> to rank *below* *H. This is because we need that PREC clause's L not to inhibit us from deleting a solitary High tone in situations where we have previously deleted another High. If the two clauses are part of single constraint, then these ranking requirements are inconsistent, but if we make them separate constraints, then the inconsistency is eliminated.

The last step here is to understand why having unsplit PREC constraints is not a problem in the analysis of mutual-counterbleeding interactions where the two processes violate different faithfulness constraints. In the hypothetical example of mutual counterbleeding in section 4.1, the PREC constraint at work is PREC(MAX-h, IDENT(voc)). Its crucial work is done by its Max-h←IDENT(voc) clause: having this ranked above *GLIDE serves to inhibit vocalization from happening in cases when vocalization will not counterbleed earlier /h/-deletion. If the PREC constraints are unsplit, then this ranking will also result in the "*IDENT(voc)-THEN-MAX-h" clause ranking above *GLIDE. This is not a problem, because in the intended winner in that mutual counterbleeding scenario, the winner did not have /h/-deletion occurring after vocalization. (Indeed, there was not any candidate with such an ordering in the candidate set depicted, and indeed there could not be given the assumptions that underlay the exposition of the hypothetical example, i.e. that deletion of /h/ before a vowel is not harmonically improving in the language in question.)

4.3 Do mutual-counterbleeding interactions exist?

As mentioned earlier, the mutual-dissimilation scenario in the previous section is based on facts in the Bantu language Tachoni (Odden 2008). Verbs in this language are underlyingly either H-toned or toneless. Tachoni also has a tense/aspect marker which consists of a (floating) melodic High tone. This marker manifests itself as the addition of an H tone to toneless verbs, but in underlyingly H-toned verbs neither the stem High nor the melodic High surfaces:

(53)	[oxu-bal-a]	'to count'	[ba-li-bála]	'they will count'
	[oxu-bék-a]	'to shave'	[ba-li-beka]	'they will shave'

Odden's (2008) rule-based analysis of these facts is that the presence of the stem H tone either causes delinking of the melodic H, or else prevents it from docking in the first place. The presence of the floating melodic H then triggers application of a rule which deletes the stem H; finally the floating H is deleted by some clean-up rule, or else it remains in place but goes phonetically uninterpreted due to its unlinked status. In this analysis, the melodic H which is initially delinked rather than outright deleted serves as a diacritic marker of derivational history. First, the melodic H is delinked, in effect marking it to be deleted later, but it remains in the representation to condition deletion of the stem H. In this way, Odden's (2008) analysis obtains the effect of mutual dissimilation without needing simultaneous application.[18]

This illustrates an important point, namely that using diacritic features to pass information about earlier stages of a derivation on to later stages is always in principle available, often allowing a particular theory of rule application (e.g. 'rules apply one at a time and can't refer to prior history') to model interactions which at first glance might seem to counter-exemplify that theory.[19] Naturally, phonologists are accustomed to assuming some sort of limits to the ability of phonological rules to introduce or manipulate diacritic features (Kiparsky 1973 *et seq.*), so this gambit—even if always available in a formal sense—will not always pass most phonologists' intuitive plausibility test. Odden's (2008) analysis of Tachoni is an example in which the posited 'diacritic'—a delinked tone which is left floating—is not especially implausible at all, given an autosegmental theory of tonal representations (Goldsmith 1976).

I know of only two works in the published literature which have explicitly advocated the existence of mutual-counterbleeding interactions.[20] The first, Ballard (1971), actually deals mainly with mutually-counterbleeding historical changes rather than synchronic rules. Moreover, the example given, of certain developments in the history of the Wu dialects of eastern China, is presented so briefly as to make it hard to judge either its plausibility or its compatibility with OT-CC. Ballard (1971) speculates that this type of interaction may also occur in synchronic grammars, but does not suggest any examples.

The second is Bliese (1975), who argues for the existence of a case of mutual dissimilation in 'Afar. The argument, however, relies crucially on the assumption that two alternations in which /a/ raises to [o] (one in which a following /a/ is part of the environment, and one in which a following [i] derived from underlying /a/ is part of the environment) can be collapsed together as a single dissimilation rule. This rule would be assumed to interact in a mutually-counterbleeding fashion with the /a/ → [i] rule. The result of collapsing together the two /a/ → [o] rules, however (p. 104) is

[18] A reviwer points out that these alternations could also be analyzed as a morphophonological exchange rule: to mark the tense/aspect morpheme, a High tone is added to stems that lack one and deleted from stems that have one. As discussed earlier in §3.2, the existence of such processes is quite controversial; however, to the extent that solid examples may be documented, it is conceivable that they could be analyzed as mutual counterbleeding along the lines given for the High-tone dissumilation scenario.

[19] For a syntactic (and fully diacritic) analogue of the analysis of Tachoni under discussion, cf. Postal's (1970) use of a feature [DOOM] to mark certain items for later deletion; see also Postal (1972: 140). For further discussion of diacritic marking as a strategy for accommodating global rules, see Kenstowicz & Kisseberth (1970), Lightner (1971: 531-532), Lehmann (1972: 542-543), Dinnsen (1974), McCawley (1975: 175) and Harris (1993: 181-182&ff.); and see Levine (1976) and Pelletier (1980) on how the possibility of such marking can erase apparent differences in the generative power of different theories of rule ordering.

[20] For a third possible example, see van Oostendorp (2008), who notes that some of the Latin facts discussed in Wells (1949) and Goldsmith (2008) appear to call for simultaneous application.

extremely cumbersome, so it does not seem that we gain a lot in terms of analytic parsimony by going with this analysis. So as with mutual counterfeeding, it would seem that—while the current literature does provide some leads—conclusive evidence for the existence of mutual counterbleeding interactions has yet to be found.

4.4 Why should we care whether mutual counterbleeding exists?

As we saw, mutual counterbleeding interactions, as analyzed in OT-CC, would actually represent covert cases of obligatory counterbleeding. In an obligatory-counterbleeding interaction, some process *Y* is blocked from applying when it would not be crucially preceded by some other process *X*. One way for there to be such a crucial ordering is if *Y* stands in a (counter)bleeding functional relationship to *X*. If this is so, then the only way for *X* and *Y* to both occur is if *X* goes first. Thus, *Y* applies only when there is a preceding instance of *X* for *Y* to counterbleed. It is not clear that such interactions exist. (Wolf 2008: §4.4.1 discusses as a possible example the interaction of assibilation and vowel shortening in Chimwi:ni [Kisseberth & Abasheikh 1975; Kentowicz & Kisseberth 1977; Hyman 1993]; I know of no possible examples besides this one.)

The other way for *Y* to be crucially preceded by *X* is for *X* to stand in a (counter)feeding functional relationship to *Y*—that is, if doing *X* creates new configurations which meet the structural description of *Y*. If *Y* is barred from occurring except when *X* crucially precedes, then *Y* is applying in environments derived by *X*, and is blocked otherwise. Unlike mutual counterbleeding, this sort of "a process is blocked except when something else crucially occurred earlier" scenario is richly attested: this is the familiar phenomenon of nonderived environment blocking (NDEB), instances of which are also referred to as derived-environment effects (DEEs).

NDEB was first brought to the attention of the field by Kenstowicz & Kisseberth (1970) and Kiparsky (1973); Wolf (2008: 247-248) provides an extensive list of possible examples of NDEB which have been proposed in the literature since then. Wolf (2008) also shows that—like obligatory counterbleeding—NDEB becomes modelable in OT-CC if we lift the ranking metaconstraint. Additionally, it is shown there that the OT-CC approach makes several restrictive and arguably correct predictions about the typology of NDEB. Most notably, the OT-CC approach to NDEB predicts that NDEBed process are barred from occurring in various kinds of vacuously-derived environments, something which many other theories of NDEB do not achieve except by stipulation.

A second motivation, besides NDEB, for lifting the ranking metaconstraint, is that this would make it possible for the PREC constraints to be ranked above the faithfulness constraints in the initial state of L1 acquisition. As first suggested by Wolf (2008: §4.4.2) and as more fully developed by Tihonova (2009), this assumption about the initial state would make it possible to model emergent, non-target-like opacity in child phonology using OT-CC. As with NDEB, examples of this are abundantly attested, suggesting again that lifting the ranking metaconstraint would be a desirable move.

There remains, though, one potential liability for the proposals that OT-CC is the appropriate framework for analyzing NDEB or emergent child opacity: a version of OT-CC shorn of the ranking metaconstraint, and therefore capable of modeling NDEB, also predicts the existence of obligatory counterbleeding, and also, as we have seen in the present paper, of mutual counterbleeding as well. Therefore, the existence (or non-existence) of obligatory counterbleeding or mutual counterbleeding interactions is a consideration of no small importance in helping us to decide whether to accept these extensions of OT-CC's empirical coverage. Alternatively, if we do accept the proposal that the ranking metaconstraint be lifted but convincing examples of obligatory or mutual counterbleeding fail to turn up, then some new assumption will have to be added to OT-CC to account for their absence.

In terms of theory comparison, it is interesting to note that the metaconstraint-less version of OT-CC is not the only theory of NDEB in OT which also predicts obligatory countebleeding. Both NDEB and obligatory counterbleeding arise from PREC constraints saying that process Y (which is otherwise generally applicable) cannot apply unless it's been preceded by some other process X. A different way of connecting the obligatoriness of different processes with one another is found in Łubowicz's (2002) proposal that NDEB arises from the local conjunction (Smolensky 1995) of markedness and faithfulness constraints. A locally conjoined constraint is created by combining together two independently-existing constraints A and B to create a new constraint $[A \& B]_D$. This combined constraint is violated only just in case constraints A and B are both violated in some domain D.

Łubowicz's (2002) idea is that derived-environment effects arise from $[M\&F]_D$ conjunction, where M is the markedness constraint that motivates the NDEBed process, and F is the faithfulness constraint that is violated by the process that creates the derived environment in which the NDEBed process applies. Normally, it is not obligatory in the language to obey markedness constraint M, but when the F-violating process occurs, then it becomes necessary (within the domain of conjunction) to obey M. Faithfulness constraint F is already violated, and so the conjoined constraint would then be violated if markedness constraint M was also violated. Informally, violation of F 'activates' the markedness constraint M which is otherwise not enforced.

This same device can also be used to produce obligatory counterbleeding, because there is no requirement that the functional relationship between the F-violating 'activating' process and the M-motivated 'activated' process is a feeding one: the conjoined constraint simply says that, within domain D, M-violation is not tolerated if an F-violating process has occurred. To illustrate this, suppose that the two processes at work are raising of word-final mid vowels (/e, o/ → [i, u]) and epenthesis of glottal stop at the end of a vowel-final word (driven by FINAL-C). An obligatory counterbleeding interaction of these processes would mean that /ʔ/-epenthesis happened only to the right of erstwhile mid vowels which had raised to high; words ending in an underlying high (or low) vowel would not undergo /ʔ/-epenthesis. That is to say, epenthesis would only occur when it would counterbleed word-final raising. The following tableau shows how we can achieve this using [M&F] conjunction:

(54)

/e#/	*MID#	[ID(hi) & FINAL-C]$_{seg}$	[DEP-C & *MID]$_{\sigma}$	IDENT (high)	DEP-C	*MID	FINAL-C
a. [e#]	W1			L	L	W1	W1
b. [i#]		W1		1	L		W1
c. [eʔ#]			W1	L	1	W1	
d. ☞ [iʔ#]				1	1		

The fully-faithful candidate [e#] is ruled out by undominated *MID#, which forbids word-final mid vowels. That markedness constraint outranks two faithfulness constraints, IDENT(high) and DEP-C, so the *MID# violation could be 'repaired' in one of two ways: raising the mid vowel to high, or epenthesizing a consonant to make the mid vowel no longer be word final.

Candidate (54b) raises the /e/ to [i], which gets rid of the violation of *MID#. However, it now violates the conjoined constraint [ID(hi) & FINAL-C]$_{seg}$. This constraint is violated by any single vowel which is both unfaithful to its input correspondent's [high] specification (which violates IDENT(high)) and which is word final (which violates FINAL-C). Informally, raising 'activates' a requirement that words cannot end in a vowel. Candidate (54c) eliminates the *MID# violation by epenthesizing a glottal stop. However, in so doing it runs afoul of a different conjoined constraint, [DEP-C & *MID]$_{\sigma}$. This constraint forbids syllables which contain both an epenthetic consonant (which violates DEP-C) and a mid vowel (which violates *MID). Informally, glottal-stop-epenthesis 'activates' a ban on mid vowels, even if not word-final. Thus, the only way to eliminate the *MID# violation without violating either conjoined constraint is to both raise the mid vowel and epenthesize the consonant, as in candidate (54d), which is our winner.

By contrast, when a word ends in a non-mid vowel, there is no glottal-stop epenthesis, because DEP-C outranks FINAL-C:

(55)

/i#/	*MID#	[ID(hi) & FINAL-C]$_{seg}$	[DEP-C & *MID]$_{\sigma}$	IDENT (high)	DEP-C	*MID	FINAL-C
☞ [i#]							1
[iʔ#]					W1		L

We thus have modeled the obligatory counterbleeding interaction between final raising and final /ʔ/-epenthesis: glottal stop epenthesis, which counterbleeds raising of underlyingly final mid vowels, only occurs when raising also happens.

Is there any reason why theories that can model NDEB effects might tend to also allow obligatory counterbleeding? As the discussion at the beginning of this sub-section attempted to bring out, there is a close conceptual link between the two notions: both involve some process, otherwise not applicable, becoming obligatory just in case some other process has occurred 'nearby' either in derivational (OT-CC) or structural (M&F conjunction) terms. This indicates that 'does this theory allow obligatory counterbleeding?' may be a question that we want to ask generally when assessing any proposal about how to model NDEB in OT.

5. Conclusion

In this paper, we have examined the properties of four different versions of OT-CC which correspond to all the possible combinations of two theoretical choices: whether or not to split the Prec constraints in two, and whether or not to retain the ranking metaconstraint. Different sets of of kinds of opaque interactions are possible under each combination of these assumptions. As a result (assuming these to be the only relevant parameters defining versions of OT-CC under consideration), we can then define the following implicational relationships between the modelability, in OT-CC, of different kinds of opaque interactions:

(56)

<table>
<tr><td colspan="2">Mutual counterbleeding where the processes violate the same BFC
(requires split Prec and lifting of ranking metaconstraint)</td></tr>
<tr><td>⇓</td><td>⇓</td></tr>
<tr><td>Mutual counterfeeding
Self-counterfeeding
(require split Prec)</td><td>NDEB (Wolf 2008: ch. 4)
Obligatory counterbleeding
(Wolf 2008: ch. 4)
Mutual counterbleeding where the processes violate different BFCs
(require lifting of ranking metaconstraint)</td></tr>
<tr><td>⇓</td><td>⇓</td></tr>
<tr><td colspan="2">'Ordinary' counterbleeding (McCarthy 2007a)
'Ordinary' counterfeeding (McCarthy 2007a)
Counterfeeding from the past (Wilson 2006)
Opaque feeding[21] (Lee 2007)
(all possible under original OT-CC proposal with unsplit Prec and ranking metaconstraint in force)</td></tr>
</table>

(BFC = basic faithfulness constraint)

We also saw that a number of possible examples of mutual counterfeeding and of mutual counterbleeding have been reported; however, probably none of them can yet be seen as totally convincing.

These findings are interesting for two reasons. First, they establish that at least certain versions of OT-CC are not unlimited in their ability to model global-rule effects. The original version of OT-CC, with unsplit Prec and the ranking metaconstraint, allows neither mutual counterfeeding nor mutual counterbleeding. If such interactions don't exist, then this is obviously a desirable prediction. In addition, should a clear example

[21] See Baković (2007a) on feeding interactions which are opaque.

of one of these interactions be found, we now know how we would need to revise the original version of OT-CC in order to bring this interaction into the scope of the theory. Second, the uncovering of implicational relationships among the types of opaque interactions that are modelable in different versions of OT-CC will be useful for future research on the theory. For example, we know that the existence or non-existence of mutual counterbleeding will bear on the question of whether we want to adopt OT-CC as our account of NDEB, but that the existence or non-existence of mutual counterfeeding will not.

For both of these reasons, these results will be important in the long run in helping to decide whether OT-CC will be viable as a theory of which forms of opacity can and cannot exist in human language. As noted in the introduction, predictions about mutual counterfeeding and mutual counterbleeding have a long history in phonological theory as arguments for and against various theories of opacity, and OT-CC in its various conceivable forms will certainly be no exception.

References

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

Anderson, Stephen R., and Wayles Browne (1973). On keeping exchange rules in Czech. *Papers in Linguistics* **6**, pp. 445-482.

Baković, Eric (2007a). A revised typology of opaque generalizations. *Phonology* **24**, pp. 217-259. [ROA-850]

Baković, Eric (2007b). Mutually-assured destruction. Post on Phonoloblog, February 19. [Available online at http://camba.ucsd.edu/blog/phonoloblog/2007/02/19/mutually-assured-destruction]

Ballard, W.L. (1971). Simultaneous ordering of phonological rules. *International Journal of American Linguistics* **37**, pp. 188-189.

Beckman, Jill N. (1998). *Positional Faithfulness*. Ph.D. dissertation, University of Massachusetts, Amherst. [ROA-234]

Bhatia, Tej, and Michael J. Kenstowicz (1972). Nasalization in Hindi: A reconsideration. *Papers in Linguistics* **5**, pp. 202-212.

Bliese, Loren F. (1975). ‘Afar vowel dissimilation: A problem in rule ordering. *Anthropological Linguistics* **17**, pp. 102-106.

Bloomfield, Leonard (1939). Menomini morphophonemics. *Travaux du cercle linguistique de Prague* **8**, pp. 108-115.

Boersma, Paul (2001). Phonology-semantics interaction in OT, and its acquisition. In Robert Kirchner, Joe Pater, and Wolf Wilkey (eds.), *Papers in Experimental and Theoretical Linguistics* **6**: *Workshop on the Lexicon in Phonetics and Phonology*. Edmonton: Department of Linguistics, University of Alberta, pp. 24-35. [ROA-369]

Bright, William (1957). *The Karok Language*. [University of California Publications in Linguistics, no. 13.] Berkeley: University of California Press.

Bromberger, Sylvain, and Morris Halle (1989). Why phonology is different. *Linguistic Inquiry* **20**, pp. 51-70.

Casali, Roderic (1997). *Resolving Hiatus*. Ph.D. dissertation, University of California, Los Angeles. [ROA-215]
Chafe, Wallace (1968). The ordering of phonological rules. *International Journal of American Linguistics* **34**, pp. 15-36.
Chomsky, Noam (1972). *Studies in Semantics in Generative Grammar*. The Hague: Mouton.
Chomsky, Noam, and Morris Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.
Coats, Herbert S. (1970). Rule environment features in phonology. *Papers in Linguistics* **2**, 110-140.
Dinnsen, Daniel A. (1974). Constraints on global rules in phonology. *Language* **50**, pp. 29-51.
Elfner, Emily (2009). Syllabification and stress-epenthesis interactions in Harmonic Serialism. Ms., University of Massachusetts Amherst. [ROA-1047]
Fought, John C. (1973). Rule ordering, interference, and free alternation in phonology. *Language* **49**, pp. 67-86.
Goldsmith, John (1976). *Autosegmental Phonology*. Ph.D. dissertation, MIT, Cambridge, MA.
Goldsmith, John (2008). Generative phonology in the late 1940s. *Phonology* **25**, pp. 1-23.
Gouskova, Maria, and Nancy Hall (2009). Acoustics of unstressable vowels in Lebanese Arabic. In Steve Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox, pp. 203-225.
Grimshaw, Jane (1997). Projection, heads, and optimality. *Linguistic Inquiry* **28**, pp. 373-422. [ROA-68]
Halle, Morris, and Valdis Zeps (1966). A survey of Latvian morphophonemics. *Quarterly Progress Report of the Research Laboratory of Electronics* **83**, pp. 104-113.
Hammond, Michael (1994). An OT account of variability in Walmatjari stress. Ms., University of Arizona, Tucson. [ROA-20]
Harris, Randy Allen (1993). *The Linguistics Wars*. Oxford: Oxford University Press.
Harris, Zellig S. (1942). Morpheme alternants in linguistic analysis. *Language* **18**, pp. 169-180.
Harris, Zellig S. (1951). *Methods in Structural Linguistics*. Chicago: University of Chicago Press.
Hill, Jane H. (1970). A peeking rule in Cupeño. *Linguistic Inquiry* **1**, pp. 534-539.
Howard, Irwin (1972). *A Directional Theory of Rule Application in Phonology*. Ph.D. dissertation, MIT, Cambridge, MA.
Hyman, Larry M. (1993). Problems for rule ordering in phonology: Two Bantu test cases. In John Goldsmith (ed.), *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago: University of Chicago Press, pp. 195-222.
Itô, Junko, and Armin Mester (2003). On the sources of opacity in OT: Coda processes in German. In Caroline Féry and Ruben van de Vijver (eds.), *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, pp. 271-303. [ROA-347]
Johnson, C. Douglas (1972). *Formal Aspects of Phonological Description*. The Hague: Mouton.
Kavitskaya, Darya, and Peter Staroverov (2010). When an interaction is both opaque and transparent: The paradox of fed counterfeeding. *Phonology* **27**, pp. 255-288.

Kenstowicz, Michael J., and Charles W. Kisseberth (1970). Rule ordering and the asymmetry hypothesis. *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS, pp. 504-519.
Kenstowicz, Michael J., and Charles W. Kisseberth (1977). *Topics in Phonological Theory*. New York: Academic Press.
Kiparsky, Paul (1973). Phonological representations. In Osamu Fujimura (ed.), *Three Dimensions of Linguistic Theory*. Tokyo: TEC, pp. 1-135.
Kiparsky, Paul (2000). Opacity and cyclicity. *The Linguistic Review* **17**, pp. 351-366.
Kiparsky, Paul (2001). Stratal OT or Sympathy? Talk presented at University of Massachusetts, Amherst. [Not seen; handout cited by McCarthy (2007a).]
Kisseberth, Charles W. (1970). On the functional unity of phonological rules. *Linguistic Inquiry* **1**, pp. 291-306.
Kisseberth, Charles (1973). Is rule ordering necessary in phonology? In Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli, and Sol Saporta (eds.), *Issues in Linguistics: Papers in Honor of Henry and Renée Kahane*. Urbana: University of Illinois Press, pp. 418-441.
Kisseberth, Charles, and Mohamed Imam Abasheikh (1975). The perfect stem in Chi-Mwi:ni and global rules. *Studies in African Linguistics* **6**, pp. 249-266.
Koutsoudas, Andreas, Gerald Sanders, and Craig Noll (1974). The application of phonological rules. *Language* **50**, pp. 1-28.
Lakoff, George (1970). Global rules. *Language* **46**, pp. 627-639.
Lakoff, George (1972). The global nature of the nuclear stress rule. *Language* **48**, pp. 285-303.
Lee, Minkyung (2007). OT-CC and feeding opacity in Javanese. *Studies in Phonetics, Phonology, and Morphology* **13**, pp. 333-350. [Available online at http://society.kisti.re.kr/~pmc/work/vol_13_2/08.doc]
Lehmann, Twila (1972). Some arguments against ordered rules. *Language* **48**, pp. 541-540.
Lightner, Theodore M. (1971). Generative phonology. In William Orr Dingwall (ed.), *A Survey of Linguistic Science*. College Park: Linguistics Program, University of Maryland, pp. 498-564.
Lightner, Theodore M. (1972). *Problems in the Theory of Phonology, vol. 1: Russian Phonology and Turkish Phonology*. Edmonton: Linguistic Research.
Łubowicz, Anna (2002). Derived environment effects in Optimality Theory. *Lingua* **112**, pp. 243-280.
McCarthy, John J. (1996). Remarks on phonological opacity in Optimality Theory. In Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky (eds.), *Studies in Afroasiatic Grammar: Papers from the Second Conference on Afroasiatic Linguistics, Sophia Antipolis, 1994*. The Hague: Holland Academic Graphics, pp. 215-243.
McCarthy, John J. (1999). Sympathy and phonological opacity. *Phonology* **16**, pp. 331-399.
McCarthy, John J. (2000). Harmonic serialism and parallelism. In Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-yung Kim (eds.), *Proceedings of the North East Linguistic Society 30*. Amherst: GLSA, pp. 501-524.
McCarthy, John J. (2003). Comparative Markedness. *Theoretical Linguistics* **29**, pp. 1-51.

McCarthy, John J. (2007). *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.
McCarthy, John J. (2010). Studying GEN. *Journal of the Phonological Society of Japan* **13**, pp. 3-12. [ROA-1049]
McCarthy, John J., and Alan Prince (1994). The Emergence of the Unmarked: Optimality in Prosodic Morphology. In Mercè González (ed.), *Proceedings of the North East Linguistic Society 24*. Amherst: GLSA, pp. 333-379. [ROA-13]
McCarthy, John J., and Alan Prince (1995). Faithfulness and reduplicative identity. In Jill N. Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey (eds.), *University of Massachusetts Occasional Papers in Linguistics* **18**: *Papers in Optimality Theory*. Amherst: GLSA, pp. 249-384. [ROA-60]
McCawley, James D. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.
McCawley, James D. (1973). Global rules and Bangubangu tone. In Michael J. Kenstowicz and Charles W. Kisseberth (eds.), *Issues in Phonological Theory*. The Hague: Mouton, pp. 160-168.
McCawley, James D. (1975). Review of *The Sound Pattern of English*. In Didier L. Goyaverts and Geoffrey K. Pullum (eds.), *Essays on* The Sound Pattern of English. Ghent: E. Story-Scientia, pp. 145-197.
Miller, D. Gary (1973). Some theoretical implications of Greenlandic phonology. *Papers in Linguistics* **6**, pp. 371-427.
Miller, D. Gary (1975). On constraining global rules in phonology. *Language* **51**, pp. 128-132.
Moreton, Elliott (1999). Non-computable functions in Optimality Theory. Ms., University of Massachusetts Amherst. [ROA-364]
Myers, James, and Jane Tsay (2002). Neutralization in Taiwanese tone sandhi. Ms., National Chung Cheng University, Min-Hsiung, Chia-Yi, Taiwan. [Available online at http://www.ccunix.ccu.edu.tw/%7Elngproc/MyersTsayTTSneut.pdf]
Narang, G.C., and Donald A. Becker (1971). Aspiration and nasalization in the generative phonology of Hindi-Urdu. *Language* **47**, pp. 646-667.
Odden, David (2008). Ordering. In Bert Vaux and Andrew Nevins (eds.), *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press, pp. 61-120.
van Oostendorp, Marc (2008). esti este! Post on Phonoloblog, June 25. [Available online at http://camba.ucsd.edu/blog/phonoloblog/2008/06/25/esti-este]
Pelletier, Francis Jeffrey (1980). The generative power of rule ordering in formal grammars. *Linguistics* **18**, pp. 17-72.
Piggott, G. L. (1975). More on the application of phonological rules. *Montreal Working Papers in Linguistics* **5**, pp. 113-150.
Postal, Paul M. (1968). *Aspects of Phonological Theory*. New York: Harper and Row.
Postal, Paul M. (1970). On coreferential complement subject deletion. *Linguistic Inquiry* **1**, pp. 439-500.
Postal, Paul M. (1972). The best theory. In Stanley Peters (ed.), *Goals of Linguistic Theory*. Englewood Cliffs, NJ: Prentice-Hall, pp. 131-170.
Press, M.L. (1979) *Chemehuevi: A Grammar and Lexicon*. [University of California Publications in Linguistics, no. 92.] Berkeley: University of California Press.

Prince, Alan (2002). Entailed ranking arguments. Ms., Rutgers University, New Brunswick, NJ. [ROA-500]
Prince, Alan, and Paul Smolensky (2004 [1993]). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell. [ROA-537]
Pulleyblank, Douglas (2006). Minimizing UG: Constraints upon constraints. In Donald Baumer, David Montero, and Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 15-39. [Lingref #1430]
Pullum, Geoffrey K. (1976). Sequential and simultaneous rule application in Spanish phonology. *Lingua* **38**, pp. 221-262.
Smith, Jennifer L. (to appear). The formal definition of the ONSET constraint and implications for Korean syllable structure. In Toni Borowsky, Shigeto Kawahara, Takahito Shinya, and Mariko Sugahara (eds.), *Prosody Matters: Essays in Honor of Lisa Selkirk*. London: Equinox. [Available online at http://www.unc.edu/~jlsmith/home/pdf/smith2010_onset_korean.pdf]
Smolensky, Paul (1995). On the structure of the constraint component CON of UG. Talk presented at University of California, Los Angeles. [Handout available as ROA-86]
Sohn, Ho-min (1975). *Woleaian Reference Grammar*. Honolulu: University of Hawai'i Press.
Steriade, Donca (2001). Directional asymmetries in place assimilation. In Elizabeth Hume and Keith Johnson (eds.), *The Role of Speech Perception in Phonology*. San Diego: Academic Press, pp. 219-250.
Tesar, Bruce (1995). *Computational Optimality Theory*. Ph.D. dissertation, University of Colorado, Boulder. [ROA-90]
Tesar, Bruce, and Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.
Tihonova, Olga (2009). *Acquisition and Opacity*. Master's thesis, Universititet i Tromsø. [ROA-1043]
Trommer, Jochen (2007). Voicing and polarity in Luo. Ms., Universität Leipzig. [Available online at http://www.uni-leipzig.de/~jtrommer/papers/antiauto3.pdf]
Underhill, Robert (1976). Noun bases in two Eskimo dialects: A study in comparative morphophonemics. In Eric P. Hamp (ed.), *Papers on Eskimo and Aleut Linguistics*. Chicago: CLS, pp. 239-271.
Vago, Robert M. (1977). In support of extrinsic ordering. *Journal of Linguistics* **13**, pp. 25-41.
Vago, Robert M., and Edwin Battistella (1982). Rule application in phonology. Ms., City University of New York.
Wells, Rulon S. (1949). Automatic alternation. *Language* **25**, pp. 99-116.
Wheeler, Max (1979). *Phonology of Catalan*. Oxford: Blackwell.
Wilbur, Ronnie Bring (1973). Reduplication and rule ordering. In Claudia Corum, T. Cedric Smith-Stark, and Ann Weiser (eds.), *Proceedings of the 9th Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS, pp. 679-687.
Wilson, Colin (2006). Counterfeeding from the past. Ms., University of California, Los Angeles. [Available online at

http://camba.ucsd.edu/blog/sadphig/files/2009/01/counterfeedingfromthepast.pdf]
Wolf, Matthew (2008). *Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model*. Ph.D. dissertation, University of Massachusetts, Amherst. [ROA-996]
Wolf, Matthew (2010). On the existence of counterfeeding from the past. Talk presented at 84th Linguistic Society of America Annual Meeting, Baltimore. [Handout available online at http://wolf.phonologist.org/OnTheExistenceOfCFFTP-corrected.pdf]
Zhang, Jie, Yuwen Lai, and Craig Turnbull-Sailor (2006). Wug-testing the "tone circle" in Taiwanese. In Donald Baumer, David Montero, and Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 453-461. [lingref #1479]