

## **UC Santa Cruz**

### **Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince**

#### **Title**

Elsewhere Effects in Optimality Theory

#### **Permalink**

<https://escholarship.org/uc/item/1m56m1ht>

#### **Author**

Bakovic, Eric

#### **Publication Date**

2006-06-20

## Elsewhere Effects in Optimality Theory\*

Eric Baković  
UC San Diego

[T]he Elsewhere Condition folds together a point of logic (PTC) with additional claims about what linguistic phenomena are incompatible. With the incompatibility claims properly factored out into substantive constraints of various types, what's left is PTC; that is to say, nothing.

— Prince & Smolensky (1993/2004)

### 1. Introduction

My goal in this paper is to demonstrate how the basic logic of constraint ranking in Optimality Theory (OT; Prince & Smolensky 1993/2004) directly predicts the disjunctive application of processes in an ‘elsewhere’ relationship without the need for a separate principle like the Elsewhere Condition (the EC; Anderson 1969, 1974, Kiparsky 1973) and its attendant problems of formulation in the theory of ordered string-rewriting rules (SPE; Chomsky & Halle 1968). The various details of the empirically correct formulation of the EC (Halle 1995, Halle & Idsardi 1997) that must be independently stipulated in SPE all fall out as a necessary consequence of constraint ranking logic in OT.

Much of what I have to say here serves to amplify and refine the following claim made by Prince & Smolensky (1993/2004) in their discussion (§7.2.1) of the “obvious affinities” between the EC and Pāṇini’s Theorem on Constraint-ranking (PTC).

PTC is merely a point of logic, but the Elsewhere Condition is thought of as a principle specific to UG, responsible for empirical results which could very well be otherwise.

This claim clarifies a fundamental distinction between PTC and the EC, one that has been subject to some basic misinterpretations in the literature (see e.g. §4.4). PTC is a *theorem*, a proven proposition following from basic assumptions about how constraints interact in OT; it is not a principle proposed on the basis of some empirical evidence or formal desiderata. PTC does not dictate how (certain) constraints *must be* ranked with respect to each other; what it does is identify some ranking conditions under which certain constraints are expected to be rendered inert, given how constraints interact in OT.<sup>1</sup> Thus, the only way that PTC could be otherwise is for con-

---

\* When I began work on this topic 10 years ago, the not-so-small matter of a dissertation got in the way and I never really recovered the momentum until now. I thank Alan Prince, John McCarthy, Greg Lamontagne, and those at a Rutgers Linguistics Lunch gathering in late 1996 for discussion of and comments on my early thoughts and observations. Alan has also continued to discuss many of these ideas with me on the pathetically rare occasions when I would try (but fail) to revive the paper. I of course take sole responsibility for any errors that may have found their way into these pages despite all this helpful input.

<sup>1</sup> Indeed, as Prince & Smolensky (1993/2004;§5.3) clarify: “The utility of the result [= the proof of PTC, reproduced in §4.4 below —EB] is that it allows the analyst to spot certain easy ranking arguments.”

straint interaction to be defined differently; there are no ‘alternative formulations’ of PTC in OT.<sup>2</sup>

By (stark) contrast, the EC *does* dictate how rules must be ordered and, furthermore, how that ordering must be interpreted. As such, it *is* a principle put forth on the basis of some empirical evidence or formal desiderata; it does not follow in any way from any more basic assumptions about how rules interact in SPE, and as such can be refined and redefined, as the EC has in fact been several times in the literature. In the remainder of this paper, I show how various refinements of the EC, made necessary by empirical considerations and defended with formal ones, derive exactly (though quite unintentionally and accidentally) the kinds of ‘elsewhere’ effects that OT already predicts.

The significance of this demonstration should not be underestimated. Halle & Idsardi (1997:344) have said of Kiparsky (1973) that it is “in [their] opinion among the most important contributions to phonology”, an opinion about which I believe there can be no serious disagreement. OT’s basic assumptions, which are motivated on independent grounds, predicts the precise behavior of ‘elsewhere’ effects. Surely, then, there can be no disagreement that OT is also among the most important contributions to phonology.

## 2. ‘Elsewhere’ effects

I take it that the phonology of a language is primarily responsible for the distribution of (elements of) speech sounds and structures: the contexts in which phonological elements are found or not found. The concept of ‘elsewhere’ is intimately tied to this view. When a distributionally-restricted phonological element is discovered, one can state its distribution in elsewhere-language: ‘Phonological element *x* is found in context *c*; elsewhere, *x* is not found’ — where ‘elsewhere’ means ‘in contexts other than *c*’.

The analytically more interesting cases, however, are ones in which *x* is found in context *c* and some other, closely-related element *y* is found elsewhere; that is, where *x* and *y* are in complementary distribution. The complementary distribution of allophones of a phoneme is a special case of this type. In most such cases, the distribution of one allophone (the *basic* one) is predictable only when juxtaposed with the independently predictable distribution of the other (the *derived* one).<sup>3</sup> The contexts in which the basic allophone is found are too varied to be collapsed in terms of natural classes, and thus the best one can say, once one has defined the context *c* in which the derived allophone occurs, is that the basic allophone is found ‘elsewhere’ — again, in contexts other than *c*.

All examples of ‘elsewhere’ effects have this same basic character, and are thus all examples of complementary distribution; allophonic complementary distribution is simply a special case. I make a distinction here between what I will call *unbounded* and *bounded* cases of complementary distribution. Unbounded cases are those that can be described relative to all possible contexts; allophonic complemen-

<sup>2</sup> *Caveat lector*: Piggott (2003:401ff) puts forth a ranking metacondition by the same name as PTC as part of his proposed massive revision of OT, highlighting the point made in the text. (See also §4.4.)

<sup>3</sup> There may of course be multiple derived allophones; I simplify here for the sake of exposition.

tary distribution is, at least prototypically, an unbounded case of complementary distribution. Bounded cases, on the other hand, are those that must be described relative to some well-defined proper subset of all possible contexts. The distinction between these two types of ‘elsewhere’ effect is important primarily because of the different issues each raises for an SPE analysis.

### 2.1. Unbounded complementary distribution

Consider the following elementary case of allophonic complementary distribution: ‘vowels are nasalized after nasals, and oral elsewhere’. According to this descriptive statement, the incompatible segmental properties ‘nasal’ and ‘oral’, are in complementary distribution in vowels. A proper subset of vowels, those found ‘after nasals’, are nasal; the complement of that set, those found ‘elsewhere’, are oral. There are at least five ways in which such a distributionally complementary situation can be characterized in SPE.

#### 2.1.1. Prior underlying specification

The most familiar sort of analysis is one in which vowels are all underlyingly oral and a Nasalization rule nasalizes the subset of vowels that occur after nasals. Under this analysis, the ‘elsewhere’ case is the epiphenomenal outcome of all vowels being underlyingly specified as oral, as shown by the schematic derivations in (1).

- (1) Elsewhere = prior underlying specification (N = nasal, O = oral)

<u>Nasalization</u> : V → [+nas] / [+nas] __			
	<u>UR</u>		<u>SR</u>
a.	/NV <sub>o</sub> /	→	[NV <sub>N</sub> ]
b.	/OV <sub>o</sub> /	→	[OV <sub>o</sub> ]

#### 2.1.2. Underspecification + feature-filling rule application

A second possible analysis is one in which vowels are underlyingly unspecified for the nasal/oral distinction and in which rules apply in a strictly feature-filling fashion. A Nasalization rule affects the subset of vowels after nasals, followed by an Oralization rule that affects the remaining, still unspecified vowels elsewhere.<sup>4</sup> The ‘elsewhere’ case is the combined result of underspecification and feature-filling rule application, which together force the rules to operate in complementary environments, as shown in (2).

- (2) Elsewhere = underspecification + feature-filling rule application

<u>Nasalization</u> : V → [+nas] / [+nas] __		(feature-filling only)	
<u>Oralization</u> : V → [-nas]		(feature-filling only)	
	<u>UR</u>		<u>SR</u>
a.	/NV/	→	[NV <sub>N</sub> ]
b.	/OV/	→	[OV <sub>o</sub> ]

Unlike the analysis in (1), we have two rules in (2). These rules could, under standard SPE assumptions about rule ordering, be ordered differently. However, the

<sup>4</sup> This presupposes that ‘oral’ and ‘nasal’ are distinct values of the same feature (i.e., [-nas] and [+nas]); assuming that there is a privative feature [nas] would simply reduce this analysis to the one in (1).

opposite order of these rules would simply render Nasalization completely inapplicable: once the vowels are made oral, they cannot undergo strictly feature-filling Nasalization.<sup>5</sup>

### 2.1.3. Complementary contexts of application

A third possible analysis involves vowels that may be either nasal or oral underlyingly. This time there are two feature-changing rules: Nasalization nasalizes all vowels after nasals, and Oralization oralizes all vowels except those after nasals.<sup>6</sup> Because both rules are feature-changing and apply in precisely complementary environments, either order will produce the same results (as indicated by double-sided arrows in the derivations).

(3) Elsewhere = complementary contexts of application

*Nasalization*:  $V \rightarrow [+nas] / [+nas] \_$  (feature-changing)  
*Oralization*:  $V \rightarrow [-nas] / \neg([+nas] \_)$  (feature-changing)

<i>UR</i>	<i>Nasalization</i>	<i>Oralization</i>	<i>SR</i>
a. $/NV_N/ \rightarrow$	$ NV_N  \Leftrightarrow$	$n/a \rightarrow$	$[NV_N]$
b. $/NV_O/ \rightarrow$	$ NV_N  \Leftrightarrow$	$n/a \rightarrow$	$[NV_N]$
c. $/OV_N/ \rightarrow$	$n/a \Leftrightarrow$	$ OV_O  \rightarrow$	$[OV_O]$
d. $/OV_O/ \rightarrow$	$n/a \Leftrightarrow$	$ OV_O  \rightarrow$	$[OV_O]$

The problem with this sort of analysis, of course, is that it simply recapitulates the complementarity of the rules' effects via their complementary contexts of application. Though there may be situations in which this may be preferable or necessary, I assume that the standard types of complementary distribution with which I am concerned in this paper do not call for this sort of direct stipulation of complementarity.

### 2.1.4. Prior rule specification

A fourth possible analysis again involves either nasal or oral vowels underlyingly. An earlier, general rule oralizes vowels in all contexts, and a later, more specific rule nasalizes (or re-nasalizes, as the case may be) the subset of vowels in nasal contexts. This time, the 'elsewhere' case is taken care of by prior rule specification, as shown in (4).

<sup>5</sup> The analyses for which Koutsoudas *et al.* (1974:8-10) invoke their Proper Inclusion Precedence Principle (the PIPP) are variants of the sort of analysis in (2): the two rules are in a mutual bleeding ("bleeding and counter-bleeding") relationship — accomplished in (2) by the rules' strictly feature-filling application — and the structural description (SD) of one rule (in this case, Nasalization) properly includes the SD of the other (Oralization). The main difference is that the PIPP, like the EC, is meant to predict the correct interaction between two rules the SDs of which stand in an inclusion relationship (see §4).

<sup>6</sup> How to state this negative environment is itself debatable; for lack of a better example, I use the *ad hoc* notation ' $\neg([+nas] \_)$ ' in (3); see Harris (1969:40) for an example of a similar type of notation.

(4) Elsewhere = prior rule specification

Oralization:  $V \rightarrow [-nas]$  (feature-changing)

Nasalization:  $V \rightarrow [+nas] / [+nas] \_$  (feature-changing)

	<u>UR</u>		<u>Oralization</u>		<u>Nasalization</u>		<u>SR</u>
a.	$/NV_N/$	$\rightarrow$	$ NV_o $	$\rightarrow$	$ NV_N $	$\rightarrow$	$[NV_N]$
b.	$/NV_o/$	$\rightarrow$	$ NV_o $	$\rightarrow$	$ NV_N $	$\rightarrow$	$[NV_N]$
c.	$/OV_N/$	$\rightarrow$	$ OV_o $	$\rightarrow$	<i>n/a</i>	$\rightarrow$	$[OV_o]$
d.	$/OV_o/$	$\rightarrow$	$ OV_o $	$\rightarrow$	<i>n/a</i>	$\rightarrow$	$[OV_o]$

We are again forced to recognize the predicted possibility of the opposite order between Oralization and Nasalization. As in the case of (2), the prediction seems harmless: Oralization would simply undo the prior effects of Nasalization, such that Nasalization simply has no effect; this analysis is equivalent to one that simply lacks the Nasalization rule altogether. But as originally discussed by Pullum (1976), many linguists have expressed a prejudice against derivations of the kind in (4)a, whereby a form undergoes a there-and-back-again sequence of changes  $A \rightarrow B \rightarrow A$ . Pullum calls these “*Duke of York derivations*, and the strategy of postulating such a derivation in order to achieve a description of some piece of linguistic data [...] *the Duke of York gambit*” (Pullum 1976:83). I return to discuss the relevance of the Duke of York gambit in §2.2.2 below.

### 2.1.5. Disjunctive rule ordering

The fifth and final possible analysis involves the same underlying representations and rules as in (4), but with the added condition that the ordering between the rules is *disjunctive*. When a rule A is disjunctively ordered before another rule B, then B may not apply to a string that A has applied to. In this case, Nasalization would be disjunctively ordered before Oralization, which is thus blocked whenever Nasalization applies.

(5) Elsewhere = disjunctive rule ordering

Nasalization:  $V \rightarrow [+nas] / [+nas] \_$  (feature-changing)

Oralization:  $V \rightarrow [-nas]$  (feature-changing)

	<u>UR</u>		<u>Nasalization</u>		<u>Oralization</u>		<u>SR</u>
a.	$/NV_N/$	$\rightarrow$	$ NV_N $	$\rightarrow$	<i>blocked</i>	$\rightarrow$	$[NV_N]$
b.	$/NV_o/$	$\rightarrow$	$ NV_N $	$\rightarrow$	<i>blocked</i>	$\rightarrow$	$[NV_N]$
c.	$/OV_N/$	$\rightarrow$	<i>n/a</i>	$\rightarrow$	$ OV_o $	$\rightarrow$	$[OV_o]$
d.	$/OV_o/$	$\rightarrow$	<i>n/a</i>	$\rightarrow$	$ OV_o $	$\rightarrow$	$[OV_o]$

Chomsky & Halle (1968:77) originally proposed that disjunctive order between rules is predictable on the basis of their abbreviability via parentheses or angled brackets; this is obviously not possible in the case of rules like Nasalization and Oralization, which cannot be abbreviated in one of these ways, at least not in any meaningful way.<sup>7</sup> This means either that disjunctive order must be stipulated on a

<sup>7</sup> Allowing subscripted angled brackets in both the structural change and the structural description of a rule might work — but the result in this case, as shown in (i) below, is clearly preposterous:

$$(i) \quad V \rightarrow \begin{matrix} \langle [+nas] \rangle_a \\ \langle [-nas] \rangle_b \end{matrix} / \begin{matrix} \langle [+nas] \_ \rangle_a \\ \langle \_ \rangle_b \end{matrix}$$

case-by-case basis (at least, for those cases that are not notationally abbreviable) or else that some other basis for predicting disjunctive order must be found — whence the EC, as discussed in §4 below.

## 2.2. Bounded complementary distribution

There are examples of complementary distribution that must be described relative to some proper subset of all possible contexts. In such bounded cases of complementary distribution, alternations between these and other contexts require that underlying representations be freely specified as in the prior rule specification analysis in (4) and the disjunctive rule application analysis in (5), not constrained as in the prior underlying specification analysis in (1) and the underspecification + feature-filling rule application analysis in (2). Given the impossibility of the analyses in (1) and (2) and the problems identified with the analysis in (4) above, the analysis in (5) is the only remaining live possibility.<sup>8</sup>

### 2.2.1. English branching foot heads

In English, for example, long and short vowels are in complementary distribution *in the heads of branching main stress feet*.<sup>9</sup> The head vowel of a branching foot is long under the following conditions: the head vowel is [–high], and the nonhead vowel is /i/ followed by another vowel in hiatus.<sup>10</sup> If a branching foot does not meet all of these conditions — i.e., if the head vowel is [+high], or if the nonhead vowel is not /i/ followed by a vowel in hiatus; in other words, ‘elsewhere’ — then the head vowel is short.

The main stress foot in English is a bimoraic trochee built at the right edge of the prosodic word, modulo extrametricality. A branching foot contains two light syllables ( $\sigma_L$ ) as in (6)a; a nonbranching foot contains a single heavy syllable ( $\sigma_H$ ), as in (6)b. (The head/stressed syllable is denoted by  $\sigma'$ ; the post-podal syllable in angled brackets denotes the possible presence of a word-final extrametrical syllable.)

- |  |   |
|--|---|
| (6) a. Branching foot                                    | b. Nonbranching foot                            |
| ... ( $\sigma'_L \sigma_L$ ) $\langle \sigma \rangle \#$ | ... ( $\sigma'_H$ ) $\langle \sigma \rangle \#$ |

The complementarity of long and short vowels in the heads of branching feet can be appreciated by comparing (*rǎdi*) $\langle$ al $\rangle$ , in which the stressed vowel is and must be long, with (*rǎdi*) $\langle$ cal $\rangle$ , in which the stressed vowel is and must be short. In both

Other abbreviatory conventions that entail disjunctive ordering (see §4) are also of no use here.

<sup>8</sup> Recall that I am assuming that the complementary contexts of application analysis in (3) is a non-starter for the examples of complementary distribution with which I am concerned in this paper.

<sup>9</sup> Myers (1987); *cf.* Burzio (1994). On the relationship between the surface quality and underlying quantity of English vowels, see also Chomsky & Halle (1968), Halle (1977), and Halle & Vergnaud (1987). For a more recent discussion of the tense/lax vowel distinction in English, see Green (2001).

<sup>10</sup> The head vowel must also not be followed by a “strong cluster” due to closed-syllable shortening in English (see also Myers 1987). Like Halle (1995), I abstract away from this additional condition.

cases, the head vowel is [-high] and the nonhead is /i/; the difference is that the /i/ is in hiatus in the case of (*rǎdi*)⟨*al*⟩, whereas it is not in hiatus in the case of (*rǎdi*)⟨*cal*⟩.

Given the bounded context of application of these processes (head vowels of branching feet as opposed to vowels in general), simple alternations serve to determine whether a vowel in English is underlyingly long or short even when it is the head vowel of a branching foot and thus subject to one of the length-neutralizing conditions above. For instance, the branching-foot head vowels in the derived words *re*(*médi*)⟨*al*⟩ and *Shake*(*spēāri*)⟨*an*⟩ are necessarily long while the corresponding (underlined) vowels in (*rémĕ*)⟨*dy*⟩ and (*Sháke*)⟨*spēāre*⟩ contrast in length. Similarly, the branching-foot head vowels in the derived words (*mýthi*)⟨*cal*⟩, *hu*(*mídi*)⟨*ty*⟩, (*týpi*)⟨*cal*⟩, and *di*(*víni*)⟨*ty*⟩ are obligatorily short while the corresponding vowels are short in (*mýth*) and (*hú*)⟨*míd*⟩ but long in (*týpe*) and *di*(*víne*). Clearly, vowel length is underlyingly contrastive in English.

Since there is a contrast between long and short vowels, there must be two feature-changing rules in an SPE analysis. The prose statements of these rules, following Halle (1995:27), are given in (7); see also Kenstowicz (1994a:218).

- (7) Lengthening (a) and Shortening (b)
- a. Lengthen the head vowel of a branching foot iff
    - i. the head vowel is [-high], and
    - ii. the nonhead vowel is /i/ followed by a vowel in hiatus.
  - b. Shorten the head vowel of a branching foot.

Lengthening and Shortening stand in the same logical relationship to each other as do the rules of Nasalization and Oralization, respectively. First, the members of each pair make contradictory and thus incompatible structural changes: the output of Lengthening is a long vowel while the output of Shortening is a short one; the output of Nasalization is a nasal vowel while the output of Oralization is an oral one. Second, the members of both pairs of rules share a necessary condition: Lengthening and Shortening both apply to head vowels of branching feet, and Nasalization and Oralization both apply to vowels. Third, this necessary condition is sufficient for one member of each pair (Shortening, Oralization) while it is not sufficient for the other (Lengthening, Nasalization).

However, the analytical options in (1) and (2) for Nasalization and Oralization cannot be used for Lengthening and Shortening, since two feature-changing rules are necessary in the latter case. For the prior underlying specification analysis in (1), all vowels would have to be underlyingly short, which is clearly wrong; for the underspecification + feature-filling rule application analysis in (2), on the other hand, all vowels would have to be underlyingly unspecified for length, another impossibility.

This leaves us with (4) and (5), the analytical options that allow underlying contrast and feature-changing rules. In the case of the prior rule specification analysis in (4), all vowels are shortened by Shortening and then the correct subset of them are lengthened or re-lengthened, as the case may be, by Lengthening, as shown in (8).<sup>11</sup>

<sup>11</sup> For the sake of simplicity in the presentation of these and subsequent derivations, I assume that the underlying representation already includes footing and extrametricality information.



## (8) Shortening precedes Lengthening

	<u>UR</u>		<u>Shortening</u>		<u>Lengthening</u>		<u>SR</u>
a.	/ (rǎdi)⟨al⟩ /	→	(rǎdi)⟨al⟩	→	(rǎdi)⟨al⟩	→	[ (rǎdi)⟨al⟩ ]
b.	/ (rǎdi)⟨al⟩ /	→	(rǎdi)⟨al⟩	→	(rǎdi)⟨al⟩	→	[ (rǎdi)⟨al⟩ ]
c.	/ (rǎdi)⟨cal⟩ /	→	(rǎdi)⟨cal⟩	→	<i>n/a</i>	→	[ (rǎdi)⟨cal⟩ ]
d.	/ (rǎdi)⟨cal⟩ /	→	(rǎdi)⟨cal⟩	→	<i>n/a</i>	→	[ (rǎdi)⟨cal⟩ ]

**2.2.2. The Duke of York gambit**

As Prince (1997a) notes, Chomsky & Halle (1968:181, 240*ff*) propose precisely this order of Shortening before Lengthening (there called ‘LAXING’ and ‘TENSING’, respectively), while Kiparsky (1982:44) derives this same order from each rule’s behavior with respect to the Lexical-Phonological cycle (Shortening is cyclic, Lengthening is post-cyclic). Nevertheless, Kenstowicz (1994a) and Halle (1995) use this example to argue for the necessity of analytical option (5), disjunctive rule ordering (by way of the EC, as detailed in §4 below). Prince (1997a:2) wonders about the reasons for this argument:

The Kenstowicz/Halle example does not in fact provide much evidence for adjoining an Elsewhere Condition to extrinsic-ordering models of the SPE type: the relationship between the two rules is already obtained by the theory of serial ordering. Special lengthening is ordered *after* General shortening, and simply undoes the general shortening in its own narrow environment. This kind of effect, by which ‘special’ imposes itself on ‘general’ through ordinary rule-application, is predicted by ordering theories like SPE: there is no obvious need to re-predict it.

In response, Halle & Idsardi (1998:1) cite Pullum (1976) (by way of Chomsky 1995) on the “unacceptability” of Duke of York derivations like those in (8)a:<sup>12</sup>

Chomsky (1995, p. 220) has noted “a linguistic expression L cannot be defined just as a pair  $(\pi, \lambda)$  formed by a convergent derivation. Rather, its derivation must be optimal, satisfying certain natural economy conditions: locality of movement, **no ‘superfluous steps’ in derivations**, and so on” [...]. Derivations such as [(8)a], which have been dubbed by Pullum (1976) the “Duke of York gambit”, have justly been singled out as unacceptable because of the superfluous steps in them.

Neither Chomsky nor Halle & Idsardi offer any evidence or argument for the claim that derivations with “superfluous steps” are suboptimal or unnatural (nor does Pullum for that matter). Chomsky & Halle (1968:63), however, offer the following conjecture, based on some unformed ideas (as Chomsky & Halle themselves openly admit) about how the competence model might be implemented in performance.<sup>13</sup>

<sup>12</sup> Halle & Idsardi (1997:344) also invoke Pullum (1976) with regard to their analysis of *r*-deletion and *r*-insertion in Eastern Massachusetts English: “objections to [the Duke of York gambit] have been raised on the grounds that [it] subverts the essential difference between rules, which reflect idiosyncratic facts of a language, and repairs, which are consequences of general structural principles obeyed by the language.” McCarthy (1999b:6) notes that Halle & Idsardi nowhere clarify their assumptions regarding rules and repairs; see Norton (2003:125-126) for an attempt to make sense of what Halle & Idsardi might mean.

<sup>13</sup> Some of Halle’s later work is less hesitant about the implementation of the competence model in performance; see the work reprinted in Halle (2002), aptly titled *From Memory to Speech and Back*.

The question of how an internalized grammar is used in performance (speech production or perception) is of course quite open. Nevertheless, it seems reasonable to suppose that the grammar should be selected in such a way as to minimize the amount of “computation” that is necessary, and that “length of derivation” is one factor in determining “complexity of computation.” Naturally, this principle must be regarded as quite tentative. We will adhere to it where a choice arises, but we have very little evidence for or against it. To find empirical evidence bearing on a principle of this degree of abstractness is not an easy matter, but the issue is important, and one should bear it in mind in a detailed investigation of phonological structure.

It is instructive to consider Chomsky’s own strong views to the contrary on this matter, beginning with Chomsky (1965) and repeated in Chomsky (1968) and Chomsky (1995) — the latter being the source of the passage quoted by Halle & Idsardi (1998).<sup>14</sup>

(9) Chomsky on competence and performance (emphasis added)

- a. “To avoid what has been a continuing misunderstanding, it is perhaps worth while to reiterate that **a generative grammar is not a model for a speaker or a hearer.** [...] When we say that a sentence has a certain derivation with respect to a particular generative grammar, **we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation.**” (Chomsky 1965:9)
- b. “[A]lthough we may describe the grammar *G* as a system of processes and rules that apply in a certain order to relate sound and meaning, **we are not entitled to take this as a description of the successive acts of a performance model such as *PM* – in fact, it would be quite absurd to do so.**” (Chomsky 1968:117)
- c. “Recall that the ordering of operations is abstract, expressing postulated properties of the language faculty of the brain, **with no temporal interpretation implied.**” (Chomsky 1995:380)

Compare also Pullum’s (1976:100) own conclusions on the potential for any putative principle to rule out Duke of York (henceforth DY) derivations:

[A] thorough investigation of the descriptive problems in which proposed Duke of York analyses are embedded does not reveal any basis for a general constraint that would prohibit the Duke of York gambit. This means that [linguists who have expressed a prejudice against DY analyses] are in the strictest sense mistaken on a methodological matter. The guiding principle they implicitly appeal to in casting suspicion on Duke of York analysis cannot and does not exist [...] there are points both for it and against it but it cannot be trusted to select correct analyses over inadequate ones. [T]he strategy of avoiding (or of adopting) the Duke of York gambit will be reasonable precisely when the result is a reasonable analysis, and unreasonable precisely when it is not.

The analytical issue that Pullum’s paper draws attention to, then, has clearly overshadowed this concluding passage, and prejudice against DY analyses has persisted. McCarthy (1999a, 2003) and Norton (2003) are the only works known to me in which the consequences of the DY gambit are put to some sort of empirical test, and both authors conclude that the prejudice is far from unfounded.

McCarthy (1999a:§7, 2003:§2.3) argues that there are no real examples of *feeding* DY derivations: derivations in which there are three rules, where the first cru-

---

<sup>14</sup> These same passages are quoted, in a slightly different context, by McCarthy (2002:9-10).

cially feeds the second and the third rule crucially undoes the effect of the first.<sup>15</sup> For example, imagine the opposite order of Shortening and Lengthening, plus an intervening Raising rule that raises all and only long vowels. Raising would affect all vowels that are lengthened by Lengthening, and it would affect only some vowels elsewhere (those that are arbitrarily specified as long in the input, by hypothesis). Shortening will then come along and render all vowels nondistinct along the long/short dimension. This is shown in (10).

(10) Lengthening precedes Raising precedes Shortening (hypothetical)

<u>UR</u>	<u>Lengthening</u>	<u>Raising</u>	<u>Shortening</u>	<u>SR</u>
a. / (rǎdi)⟨al⟩ / →	(rǎdi)⟨al⟩   →	(rědi)⟨al⟩   →	(rědi)⟨al⟩   →	[ (rědi)⟨al⟩ ]
b. / (rǎdi)⟨al⟩ / →	(rǎdi)⟨al⟩   →	(rědi)⟨al⟩   →	(rědi)⟨al⟩   →	[ (rědi)⟨al⟩ ]
c. / (rǎdi)⟨cal⟩ / →	n/a →	(rědi)⟨cal⟩   →	(rědi)⟨cal⟩   →	[ (rědi)⟨cal⟩ ]
d. / (rǎdi)⟨cal⟩ / →	n/a →	n/a →	(rǎdi)⟨cal⟩   →	[ (rǎdi)⟨cal⟩ ]

The unattestedness of feeding DY derivations like (10) and others discussed by McCarthy is unexpected in a theory that allows DY derivations in general, thus offering some empirical support for the view that DY derivations should be disallowed.<sup>16</sup>

For his part, Norton (2003:§3.4) examines several apparent and potential cases of DY derivations and concludes “that a disjunctive interaction between contrary processes is a universal property of language” — in other words, rules that could undo each other’s effects never apply to same element. This universal property is stated as follows:

(11) Universal interaction of mutually contrary operations (Norton 2003:135)

Mutually contrary operations on segmental structure apply in distinct contexts.

This is the property of Universal Grammar (UG) that is meant to be captured by the definitions and redefinitions of the EC in the literature, ultimately resulting in failure, as I discuss in §4, due to the *ad hoc* character of the EC. Disjunctive application in the case of rules in an ‘elsewhere’ relationship is desirable, but preferably follows from independent principles of UG. As I show in §3 and then again in §4.4, this is true in OT.

### 3. OT analysis

In this section I provide OT analyses of both types of complementary distribution discussed in §2 above, demonstrating that the correct, disjunctive application of the two processes in each case is an automatic consequence of the analysis. Moreover, each analysis parallels the other in fundamental respects that reflects their essential sameness, in a way that their differing analyses in SPE never have.

---

<sup>15</sup> Derivations like those in (4) and (8), where there is no feeding of a second of three rules, are what McCarthy calls *vacuous* DY derivations. McCarthy (2003:30-31) shows that *bleeding* DY derivations (in which one rule bleeds a second, and a third undoes the first) may be reduced to the vacuous type.

<sup>16</sup> Note that feeding DY derivations technically involve no “superfluous” steps: the first step is necessary to feed the second, and the third is necessary to undo the first. Thus, disallowing DY derivations based exclusively on their superfluity incorrectly allows the unattested feeding type! (See §4.3 below.)

### 3.1. Unbounded complementary distribution

The analysis of unbounded complementary distribution in terms of the interaction of markedness and faithfulness constraints in OT has been presented and discussed at length in the literature; the particular example of nasal/oral vowel distribution from §2.1 is taken directly from the description and analysis of Madurese nasal harmony presented in McCarthy & Prince (1995:§3.3, 1999:§4.1), and that analysis is summarized here.

The OT analysis requires no constraints on underlying representations, as required in the prior underlying specification analysis in (1) and the underspecification + feature-filling rule application analysis in (2), a point that becomes particularly crucial in the analysis of bounded complementary distribution (see §3.2).<sup>17</sup> The complementarity of contexts of application needn't be stipulated as in (3), and because there is no serial derivation, there can be no Duke of York derivation as in (4). The OT analysis comes closest to being like the disjunctive rule application analysis in (5), but without the need to stipulate disjunctive application — or even to attempt to derive it from *ad hoc* principles.

In the SPE analyses reviewed in §2.1, it is assumed that vowel nasalization is the context-specific case ('after nasals') and that vowel oralization is the context-free case ('elsewhere'). This assumption is validated by universal markedness considerations: the presence of nasal vowels in the surface inventory of a language entails the presence of oral vowels in the surface inventory of that language as well. In OT, universal markedness relations like this can be formally expressed by appealing to universally fixed subhierarchies of markedness constraints (Prince & Smolensky 1993/2004); in this case (12).<sup>18</sup>

(12) Fixed markedness subhierarchy

$$*V_N \gg *V_O$$

The universally-fixed ranking in (12) states that, all else being equal, nasal vowels are more marked than oral ones because an instance of the former incurs a higher-ranked constraint violation than an instance of the latter. By ranking the conflicting faithfulness constraint IDENT(nas) in (13) with respect to the subhierarchy in (12), we describe two possible inventories. The first is described by the ranking in (14)a, an inventory with both oral and nasal vowels, each surfacing faithfully from the input due to the dominance of IDENT(nas). The ranking in (14)b describes the second inventory, with oral vowels only.<sup>19</sup>

---

<sup>17</sup> This conforms to OT's *Richness of the Base* hypothesis, that there are no language-particular restrictions on input forms (itself a corollary of the hypothesis that the only difference between grammars is the ranking of a universal constraint set). See McCarthy (2002:76-82) for discussion and references.

<sup>18</sup> The evidence for two-part fixed markedness subhierarchies like (12) is somewhat lacking; the same result (but perhaps with other consequences) can be achieved by positing that the constraint against the least-marked element (in this case,  $*V_O$ ) simply does not exist, and the same reasoning applies to the lowest-ranked constraint of any fixed markedness subhierarchy (Gouskova 2003). See also Prince (1997c, 1999, 2001) and de Lacy (2002, 2006) for a different take on universal markedness relations in general.

<sup>19</sup> Either ranking between  $*V_O$  and IDENT(nas) will do in (14)b, because any attempt to avoid an oral vowel by violating IDENT(nas) results in a nasal vowel, which violates higher-

## (13) IDENT(nas)

Corresponding segments agree in their value for the feature [ $\pm$ nas].

## (14) a. Oral and nasal vowels

IDENT(nas)  $\gg$  \*V<sub>N</sub>  $\gg$  \*V<sub>O</sub>

## b. Oral vowels, no nasal vowels

\*V<sub>N</sub>  $\gg$  { IDENT(nas), \*V<sub>O</sub> }

More precisely, the ranking in (14)b describes a context-free oralization process: any underlyingly nasal vowel is rendered oral in the optimal output candidate, violating \*V<sub>O</sub> and IDENT(nas) but sparing a violation of \*V<sub>N</sub>. To describe the complementary distribution pattern of interest, with both context-specific nasalization and context-free oralization, all that is left is the addition of the context-specific nasalization constraint in (15), \*NV<sub>O</sub>, to the top of the context-free oralization ranking in (14)b. Since the output of the nasalization process violates \*V<sub>N</sub>, \*NV<sub>O</sub> must dominate \*V<sub>N</sub>. Since it also violates IDENT(nas) when the surface nasal vowel is underlyingly oral, \*NV<sub>O</sub> must also be ranked above IDENT(nas), which it necessarily does by transitivity through \*V<sub>N</sub>.

(15) \*NV<sub>O</sub>

No [-nas] vowels following [+nas] segments.

There are other input-output mappings one could imagine that satisfy \*NV<sub>O</sub> without nasalizing the vowel and violating \*V<sub>N</sub>. Oralization of the input nasal segment, for instance, or deletion of either the nasal segment or the oral vowel, will serve to satisfy both \*NV<sub>O</sub> and \*V<sub>N</sub>. Each of these output candidates must thus violate some constraint that is ranked higher than \*V<sub>N</sub>, the highest-ranked constraint violated by the grammatical vowel nasalization output. I refer to these necessarily higher-ranked constraints collectively as OTHER, and I include one candidate that violates one of these constraints, by oralization of an input nasal segment to the left of a vowel, in the relevant tableaux.<sup>20</sup> The complete complementary distribution hierarchy is thus as in (16).

## (16) Complete complementary distribution hierarchy

{ \*NV<sub>O</sub>, OTHER }  $\gg$  \*V<sub>N</sub>  $\gg$  { IDENT(nas), \*V<sub>O</sub> }

The tableaux in (17) through (20) show the mapping that this hierarchy predicts for each of the conceivable and relevant inputs: NV<sub>N</sub>, NV<sub>O</sub>, OV<sub>N</sub>, and OV<sub>O</sub>.

## (17) Vacuous nasalization

/NV <sub>N</sub> / → [NV <sub>N</sub> ]	*NV <sub>O</sub>	OTHER	*V <sub>N</sub>	IDENT(nas)	*V <sub>O</sub>
a. ~ NV <sub>O</sub>	W		L	W	W
b. ~ OV <sub>O</sub>		W	L	W	W

ranked \*V<sub>N</sub>. Oral vowels can thus never become nasal under pressure from \*V<sub>O</sub> alone, given the fixed markedness hierarchy in (12).

<sup>20</sup> Since vowels are not contrastively nasal due to the rank of \*V<sub>N</sub> over IDENT(nas), only consonants can initiate the propagation of [ $\pm$ nas] demanded by \*NV<sub>O</sub>. Therefore, oralization of an input nasal ultimately involves oralization of a nasal consonant. But consonants do not differ in terms of [ $\pm$ nas] alone, so oralization of a consonant violates not only IDENT(nas) but also other faithfulness constraints against changing features like [ $\pm$ son]. Any such faithfulness constraint could thus be included in OTHER.

## (18) Nonvacuous nasalization

$/NV_o/ \rightarrow [NV_N]$	$*NV_o$	OTHER	$*V_N$	IDENT(nas)	$*V_o$
a. $\sim NV_o$	W		L	L	W
b. $\sim OV_o$		W	L		W

## (19) Nonvacuous oralization

$/OV_N/ \rightarrow [OV_o]$	$*NV_o$	OTHER	$*V_N$	IDENT(nas)	$*V_o$
$\sim NV_N$			W	L	L

## (20) Vacuous oralization

$/OV_o/ \rightarrow [OV_o]$	$*NV_o$	OTHER	$*V_N$	IDENT(nas)	$*V_o$
$\sim NV_N$			W	W	L

Note from these last two tableaux that the ‘elsewhere’ effect is simply due to the irrelevance of  $*NV_o$  in the already nonnasal environment to the left of the vowel. Violation of OTHER needn’t be considered, as it must be in the first two tableaux, and the more general constraint  $*V_N$  is free to rule the day in favor of the completely oral candidate.

We are now in a position to examine what the typological consequences of this analysis are. Unlike the SPE analyses in (1) and (2), underlying representations are not constrained in any way; the elsewhere result is guaranteed by the higher rank of the more specific  $*NV_o$  constraint, similar to the later ordering of the more specific Nasalization rule in the analysis in (4). A key difference between this OT analysis and the SPE analysis in (4) is that reranking of the constraints  $*NV_o$  and  $*V_N$ , unlike the reordering of Nasalization and Oralization in (4), completely masks the potential effects of the more specific  $*NV_o$  constraint because the more general  $*V_N$  constraint consistently overrides it. As the tableaux in (21) through (24) show, no candidate with a nasal vowel can ever win under this ranking of the constraints. (This is precisely the point of PTC; see §4.4.)

## (21) Nonvacuous oralization

$/NV_N/ \rightarrow [NV_o]$	$*V_N$	OTHER	$*NV_o$	IDENT(nas)	$*V_o$
a. $\sim NV_N$	W		L	L	L
b. $\sim OV_o$		W	L	W	

## (22) Vacuous oralization

$/NV_o/ \rightarrow [NV_o]$	$*V_N$	OTHER	$*NV_o$	IDENT(nas)	$*V_o$
a. $\sim NV_N$	W		L	W	L
b. $\sim OV_o$		W	L	W	

## (23) Nonvacuous oralization

/OV <sub>N</sub> / → [OV <sub>O</sub> ]	*V <sub>N</sub> ∷ OTHER	*NV <sub>O</sub>	IDENT(nas) ∷ *V <sub>O</sub>
~ OV <sub>N</sub>	W ∷		L ∷ L

## (24) Vacuous oralization

/OV <sub>O</sub> / → [OV <sub>O</sub> ]	*V <sub>N</sub> ∷ OTHER	*NV <sub>O</sub>	IDENT(nas) ∷ *V <sub>O</sub>
~ NV <sub>N</sub>	W ∷		W ∷ L

There are more constraints in the OT analysis than just \*NV<sub>O</sub> and \*V<sub>N</sub>, however, so other possible constraint rerankings must be taken into consideration. These other constraints are \*V<sub>O</sub>, IDENT(nas), and OTHER. As noted earlier, ranking \*V<sub>O</sub> above \*V<sub>N</sub> would undermine the basic assumption that nasal vowels are more marked than oral vowels, an assumption that is also implicitly made in the SPE analyses; the ranking of \*V<sub>N</sub> above \*V<sub>O</sub> must thus be held fixed for any real comparison to be made. Similarly, alternative rankings of IDENT(nas) and OTHER would define different input-output mappings for one or both processes, rendering them distinct from the structural changes made by the rules in the SPE analyses. The position of these constraints in the hierarchy must thus also be held fixed to make the right comparison between the two analyses.

In sum, then, the OT analysis of unbounded complementary distribution directly achieves the complementarity of the processes involved with nothing other than its most basic assumptions: a particular ranking of markedness and faithfulness constraints. There is no need for any artificial restrictions on underlying representations, and no need for any special conditions on ranking to avoid a Duke of York derivation. I now turn to the analysis of bounded complementary distribution, which has these same advantages.

### 3.2. Bounded complementary distribution

Consider now an OT analysis of the English lengthening and shortening facts discussed in §2.2. Just as with nasal and oral vowels, long vowels are cross-linguistically marked in comparison to short vowels: the presence of long vowels in a language's inventory entails the presence of short vowels as well. There must thus be a markedness subhierarchy stating that long vowels incur higher-ranked constraint violations than short vowels.

## (25) Fixed markedness subhierarchy (L = long, S = short)

$$*V_L \gg *V_S$$

With this subhierarchy and the conflicting faithfulness constraint IDENT(long) in (26), we describe two possible inventories.<sup>21</sup> The first is described by the ranking in

<sup>21</sup> With IDENT(long) I don't mean to suggest that there is a feature [±long], nor even that there is necessarily a single faithfulness constraint regulating both lengthening and shortening. For example, this constraint could be replaced by two separate constraints like MAX- $\mu$  and DEP- $\mu$ ; this would have some (basically trivial) consequences for the exact ranking of the constraints. (It is worth noting that the same considerations would apply if there were separate constraints like MAX-[nas] and DEP-[nas] in §2.1.)

(27)a, which is an inventory with both short and long vowels — like the inventory of English. The ranking in (27)b describes the second, an inventory with short vowels only.

(26) IDENT(long)

Corresponding vowels agree in length.

(27) a. Short and long vowels

IDENT(long)  $\gg$  \*V<sub>L</sub>  $\gg$  \*V<sub>s</sub>

b. Short vowels, no long vowels

\*V<sub>L</sub>  $\gg$  { IDENT(long), \*V<sub>s</sub> }

There must be at least two other context-sensitive markedness constraints, both ranked above the hierarchy in (27)a, that are responsible for the lengthening and shortening processes. These constraints, given in (28) and (29), are simply stated to be violated by the same strings that the Lengthening and Shortening rules in the SPE analysis apply to nonvacuously; ranking both constraints above IDENT(long) expresses the preference for the nonvacuous structural changes that the rules make.

(28) \*[-high]<sub>s</sub> / (  $\overset{\sim}{\_}$  Ci)V

No short vowel in the head of a branching foot iff

a. the head vowel is [-high] and

b. the nonhead vowel is /i/ followed by a vowel in hiatus.

(29) \*V<sub>L</sub> / (  $\overset{\sim}{\_}$   $\sigma$  )

No long vowel in the head of a branching foot.

I state these constraints in terms that most straightforwardly line up with the statements of the corresponding rules in the SPE analysis in §2.2 to facilitate comparison, though it should be obvious that neither constraint is really much of a candidate for universality. The constraint \*[-high]<sub>s</sub> / (  $\overset{\sim}{\_}$  Ci)V in (28), responsible for lengthening, is particularly complex and should be understood as an interaction among several constraints (*cf.* van de Vijver 2003). For example, the condition that the nonhead vowel be /i/ in hiatus might be governed by ONSET, which could be ranked in such a way that it can force gliding of /i/, the optimal result thus being e.g. (rād)⟨ial⟩ rather than (rādi)⟨al⟩. This analysis entails, of course, that the /i/ is in fact not the nonhead of a branching foot; instead, there is a monosyllabic foot and lengthening of its head vowel is thus a response to FTBIN.<sup>22</sup> The constraint \*V<sub>L</sub> / (  $\overset{\sim}{\_}$   $\sigma$  ) in (29), responsible for shortening, is a little more reasonable, corresponding loosely to constraints penalizing heavy-light feet; examples from the OT literature include RHHRM (Prince & Smolensky 1993/2004), NO-(HL) (Cohn & McCarthy 1994), FTHARM (Baković 1996), and GRPHARM (McCarthy 2003).<sup>23</sup>

These two constraints stand in the same logical relationship to each other as \*NV<sub>o</sub> and \*V<sub>N</sub> do. \*V<sub>L</sub> / (  $\overset{\sim}{\_}$   $\sigma$  ) penalizes long vowels in the heads of branching feet, while \*[-high]<sub>s</sub> / (  $\overset{\sim}{\_}$  Ci)V penalizes short vowels in the heads of branching feet that meet some additional conditions. Through their mutual rank above IDENT(long), these constraints make conflicting demands on the subset of branching feet that meet

<sup>22</sup> Why the head of the foot must be [-high] (e.g., (trívi)⟨al⟩, \*(trív)⟨ial⟩) and why the (would-be) nonhead of the foot cannot be /u/ (e.g., (mānu)⟨al⟩, \*(mān)⟨ual⟩) are further questions to be addressed.

<sup>23</sup> See also Prince (1990), Burzio (1994), Kenstowicz (1994b), Mester (1994), and Green (1996).



the additional conditions of the more specific constraint  $*[-\text{high}]_s / (\_Ci)V$ . For this constraint to emerge victorious, it must also dominate the more general constraint  $*V_L / (\_σ)$ .<sup>24</sup>

(30) Complementary distribution hierarchy

$$*[-\text{high}]_s / (\_Ci)V \gg *V_L / (\_σ) \gg \text{IDENT}(\text{long}) \gg *V_L \gg *V_s$$

I demonstrate the necessity of the ranking in (30) in the following tableaux. The minimally contrasting pair of words  $(\acute{r}\acute{a}di)\langle cal \rangle$  and  $(\check{r}\acute{a}di)\langle al \rangle$  are used. Each pair of tableaux shows how the correct choice is made between the two relevant output candidates for each word: the first with a long vowel in the head of the branching foot,  $(\acute{r}\acute{a}di)$ , and the second with a short vowel in that position,  $(\check{r}\acute{a}di)$ .<sup>25</sup> The tableaux come in pairs because the underlying form of each of these words could have a long vowel or a short vowel; which is correct cannot be determined on the basis of these examples alone.

Examine the shortening process first in (31) and (32). The choice between the candidates in (31) is the crucial one here; if  $(\acute{r}\acute{a}di)\langle cal \rangle$  has a long vowel underlyingly, then  $*V_L / (\_σ)$  must be ranked above  $\text{IDENT}(\text{long})$ , forcing the latter's violation in the optimal form  $(\check{r}\acute{a}di)\langle cal \rangle$ . The alternation between words like  $(\acute{t}\acute{y}pi)\langle cal \rangle$  and  $(\check{t}\acute{y}pi)\langle cal \rangle$  is thus accounted for by this ranking; the vowel is known to be underlyingly long in this case. So what if the vowel in  $(\acute{r}\acute{a}di)\langle cal \rangle$  is underlyingly short, as in (32)? This case is shown by the simple choice between the candidates in that tableau: all of the constraints, save for the lowest-ranked  $*V_s$ , favor the correct form  $(\acute{r}\acute{a}di)\langle cal \rangle$ . The lack of alternation between words like  $(\acute{m}\acute{y}th)$ ,  $(\acute{m}\acute{y}thi)\langle cal \rangle$  is thus also accounted for by this ranking.

(31) Nonvacuous shortening

$(\acute{r}\acute{a}di)\langle cal \rangle \rightarrow (\check{r}\acute{a}di)\langle cal \rangle$	$*[-\text{high}]_s / (\_Ci)V$	$*V_L / (\_σ)$	$\text{IDENT}(\text{long})$	$*V_L$	$*V_s$
~ $(\acute{r}\acute{a}di)\langle cal \rangle$		W	L	W	L

(32) Vacuous shortening

$(\acute{r}\acute{a}di)\langle cal \rangle \rightarrow (\check{r}\acute{a}di)\langle cal \rangle$	$*[-\text{high}]_s / (\_Ci)V$	$*V_L / (\_σ)$	$\text{IDENT}(\text{long})$	$*V_L$	$*V_s$
~ $(\acute{r}\acute{a}di)\langle cal \rangle$		W	W	W	L

Now let's take a look at the lengthening process, those cases where the more specific markedness constraint  $*[-\text{high}]_s / (\_Ci)V$  is relevant. Regardless of the length of the relevant vowel in the input,  $*[-\text{high}]_s / (\_Ci)V$  must be ranked above  $*V_L / (\_σ)$  to force the violation of the latter in the optimal form  $(\acute{r}\acute{a}di)\langle al \rangle$  in both (33) and (34). In (34), where the head vowel is assumed to be underlyingly short,

<sup>24</sup> For discussion of this particular ranking argument, see Burzio (1996).

<sup>25</sup> Constraints ranked higher than those considered here are assumed to narrow the candidate set down to these appropriately-footed ones; this is equivalent in all relevant respects to the assumption that Lengthening and Shortening must follow footing in the SPE analysis. Inputs are correspondingly shown here with feet, just as the underlying representations are in the SPE analysis.

IDENT(long) — which must be ranked below  $*[-\text{high}]_s / (\acute{C}i)V$  by transitivity through  $*V_L / (\acute{\sigma})$  — is also necessarily (and predictably) violated. This result accounts for the alternation between words like *(rémě)⟨dy⟩*, *re(mědi)⟨al⟩*, where the vowel is known to be underlyingly short.

## (33) Vacuous lengthening

$(\acute{r}\acute{a}di)\langle al \rangle \rightarrow (\acute{r}\acute{a}di)\langle al \rangle$	$*[-\text{high}]_s / (\acute{C}i)V$	$*V_L / (\acute{\sigma})$	IDENT (long)	$*V_L$	$*V_s$
~ $(\acute{r}\acute{a}di)\langle al \rangle$	W	L	W	L	W

## (34) Nonvacuous lengthening

$(\acute{r}\acute{a}di)\langle al \rangle \rightarrow (\acute{r}\acute{a}di)\langle al \rangle$	$*[-\text{high}]_s / (\acute{C}i)V$	$*V_L / (\acute{\sigma})$	IDENT (long)	$*V_L$	$*V_s$
~ $(\acute{r}\acute{a}di)\langle al \rangle$	W	L	L	L	W

The ranking arguments are not quite complete, however. Any other constraints that could be violated in order to satisfy the demands of the markedness constraints  $*[-\text{high}]_s / (\acute{C}i)V$  and  $*V_L / (\acute{\sigma})$  must also dominate IDENT(long), otherwise the preference for vowel-length change in both cases would not be guaranteed. One thing that these markedness constraints have in common is the ‘head of a branching foot’ specification, so some of these constraints must be responsible for footing, and these footing constraints must be assumed to be higher-ranked in any case (see footnote 25). In fact, given the relative simplicity of  $*V_L / (\acute{\sigma})$ , footing constraints are the only ones aside from IDENT(long) that can reasonably be violated to satisfy  $*V_L / (\acute{\sigma})$ .

The only other conflicting constraints left once these footing constraints are factored out, then, are constraints that can be violated to satisfy  $*[-\text{high}]_s / (\acute{C}i)V$  but not  $*V_L / (\acute{\sigma})$ . Given the relative complexity of the more specific  $*[-\text{high}]_s / (\acute{C}i)V$ , there are more than a few changes that can be performed to satisfy it. One of the conditions on  $*[-\text{high}]_s / (\acute{C}i)V$  is that the head vowel be  $[-\text{high}]$ ; faithfulness to this feature could be sacrificed. The other condition is that the nonhead vowel is  $/i/$  followed by a vowel in hiatus; this is really a complex of conditions that can be mooted by changing the nonhead vowel to something other than  $/i/$  or by inserting a consonant or deleting a vowel to break the hiatus. Since the output of the lengthening process violates  $*V_L / (\acute{\sigma})$  and IDENT(long), constraints violated by these other changes must be higher-ranked.

The opposite ranking of the constraints  $*[-\text{high}]_s / (\acute{C}i)V$  and  $*V_L / (\acute{\sigma})$  has no effect on the outcome of *(rǎdi)⟨cal⟩*; however, the outcome of *(rǎdi)⟨al⟩* is different: the more general and higher-ranked  $*V_L / (\acute{\sigma})$  would decide in favor of *(rǎdi)⟨al⟩*, with a short vowel. Indeed, a candidate satisfying  $*[-\text{high}]_s / (\acute{C}i)V$  can never win under this ranking, because the more general and higher-ranked  $*V_L / (\acute{\sigma})$  eliminates all of the relevant candidates before the more specific and lower-ranked  $*[-\text{high}]_s / (\acute{C}i)V$  gets a chance at the pickings. This means that no ranking argument can be made between  $*[-\text{high}]_s / (\acute{C}i)V$  and IDENT(long), and it is equivalent to saying in the SPE analysis that Lengthening doesn’t exist. This is shown in the tableaux in (35) and (36).

## (35) Nonvacuous shortening (hypothetical)

$(\acute{r}\acute{a}di)\langle al \rangle \rightarrow (\acute{r}\acute{a}di)\langle al \rangle$	$*V_L$ / $(\acute{\_}\sigma)$	$*[-high]_s$ / $(\acute{\_}Ci)V$	IDENT (long)	$*V_L$	$*V_s$
$\sim (\acute{r}\acute{a}di)\langle al \rangle$	W	L	L	W	L

## (36) Vacuous shortening (hypothetical)

$(\acute{r}\acute{a}di)\langle al \rangle \rightarrow (\acute{r}\acute{a}di)\langle al \rangle$	$*V_L$ / $(\acute{\_}\sigma)$	$*[-high]_s$ / $(\acute{\_}Ci)V$	IDENT (long)	$*V_L$	$*V_s$
$\sim (\acute{r}\acute{a}di)\langle al \rangle$	W	L	W	W	L

Once again, the OT analysis accounts for the complementarity of these processes without constraints on underlying representation and with no need to avoid a Duke of York derivation. The priority of lengthening over shortening is due to the only possible ranking between the more general constraint  $*V_L / (\acute{\_}\sigma)$  and the more specific constraint  $*[-high]_s / (\acute{\_}Ci)V$  that makes them both active in the grammar, because specific partially eclipses general when specific is ranked higher but general totally eclipses specific when specific is ranked lower. The defining characteristics of OT, constraint ranking and violation, are thus perfectly suited to account for both types of ‘elsewhere’ effects.

**3.3. Summary**

The primary difference between unbounded and bounded complementary distribution is the basic contrastive status of the values of the feature in question: in unbounded cases those values are noncontrastive, and in bounded cases they are (otherwise) contrastive. In the foregoing OT analyses, this is reflected in the difference between the basic rankings in (14)b vs. (27)a, repeated here as (37)a and (37)b, respectively.

- (37) a. Oral vowels, no nasal vowels                      b. Short and long vowels  
 $*V_N \gg \{ \text{IDENT}(\text{nas}), *V_o \}$                        $\text{IDENT}(\text{long}) \gg *V_L \gg *V_s$

Because the general markedness constraint  $*V_N$  is dominant in the unbounded case of (37)a, all that is necessary to ensure complementarity of the noncontrastive  $[\pm\text{nas}]$  feature is a more specific markedness constraint that conflicts with and dominates  $*V_N$ ; that constraint is  $*NV_o$ . In the bounded case of (37)b, on the other hand, the faithfulness constraint  $\text{IDENT}(\text{long})$  must be dominant in order to ensure the basic contrastiveness of vowel length. The complementarity of vowel length in certain contexts must thus be due to the interaction of two conflicting markedness constraints other than  $*V_L$  and  $*V_s$ , one more specific and higher ranked than the other; these constraints are  $*[-high]_s / (\acute{\_}Ci)V$  and  $*V_L / (\acute{\_}\sigma)$ , respectively. Focussing attention on the markedness constraints that must outrank faithfulness in both cases, then, we have the following crucial rankings.

- (38) a. Unbounded nasal, else oral                      b. Bounded long, else short  
 $*NV_o \gg *V_N$                        $*[-high]_s / (\acute{\_}Ci)V \gg *V_L / (\acute{\_}\sigma)$

This is how the essential sameness of the two types of complementary distribution is captured in the OT analyses. In both cases, the more specific constraint domi-

nates the more general, forcing violation of the more general constraint in just those specific contexts where both constraints are relevant. As demonstrated in (21)–(24) (for the unbounded case) and (35)–(36) (for the bounded case), the opposite ranking of these two markedness constraints simply renders the more specific one completely ineffectual, as if it did not even exist in the grammar. As I discuss in more detail in §4.4, this is precisely the matter of constraint ranking logic that constitutes the proof of PTC.

#### 4. Disjunctive rule application and the Elsewhere Condition

The Elsewhere Condition's *raison d'être* is (arguably) the insufficiency of Chomsky & Halle's (1968) conditions on disjunctive ordering vs. conjunctive ordering of rules.

As far as we know, the only cases of disjunctive ordering are those in which rules can be simplified in terms of parentheses and angled brackets, and in all such cases the rules are disjunctively ordered. If this is correct, we can tentatively propose the following quite strong empirical hypothesis: *where parentheses or angled brackets are required [...] for the abbreviation of a sequence of rules, these rules are disjunctively ordered; in all other cases, rules are conjunctively ordered [...]*. (Chomsky & Halle 1968:77, emphasis in the original; see also pp. 392ff.)

Several authors have proposed that certain notational conventions beyond parentheses and angled brackets also abbreviate disjunctively ordered sets of rules. Chomsky & Halle (1968:357) themselves propose that rules involving Greek-letter variables abbreviate a disjunctively ordered set, accommodating exchange rules (on which now see Anderson & Browne 1973, Alderete 1999, and Moreton 1996/1999/2004). Anderson (1969:122ff, 1974:110ff) proposes that pairs of rules abbreviable via the neighborhood (or mirror-image) convention (Bach 1968, Langacker 1969) are ordered disjunctively, and Howard (1972:92ff) proposes that the subscript notation (equivalent in most respects to the parenthesis notation; see Johnson 1972:114) abbreviates disjunctively ordered rules.

Discussion of the insufficiency of these conditions specifically as they pertain to the now-familiar 'elsewhere' cases begins with Anderson (1969:139ff, 1974:102ff), who discusses the analysis of some facts in Middle English. These facts coincidentally involve the precursor to the Modern English shortening process discussed in §2.2 and §3.2 above and its interaction with a combined open-syllable lengthening + (stepwise) lowering process. Anderson states these two rules in SPE linear form as follows.<sup>26</sup>

---

<sup>26</sup> Kiparsky (1973:94, fn. 4) notes that he had been made aware of "some problems with the Middle English example which Anderson cites"; Anderson (1974) presents the very same example in the very same way, even though he refers on p. 107 to (what may be a prepublication version of) Kiparsky (1973). I am unaware of what these problems might be; regardless, they are not germane to the point here.

(39) Middle English rules (Anderson 1969:139-140, 1974:102-103)

- a.  $\left[ \begin{array}{l} +\text{syll} \\ +\text{stress} \end{array} \right] \rightarrow [-\text{long}] / \_$  Trisyllabic Shortening (TSS)  
 — e.g. *sōri*, *sōriā* ‘sorry (sg., pl.)’  
 CVC<sub>0</sub>V
- b.  $\left[ \begin{array}{l} +\text{syll} \\ +\text{stress} \\ \langle -\text{high} \rangle \end{array} \right] \rightarrow \left[ \begin{array}{l} -\text{long} \\ -\text{high} \\ \langle +\text{low} \rangle \end{array} \right] / \_ \text{ CV}$  Open-Syllable Lengthening (OSL)  
 — e.g. *wík*, *wékes* ‘week (sg., pl.)’

When the structural descriptions of both rules are met, only TSS applies, blocking both the lengthening and lowering subparts of OSL: *běsi*, *běsines* ‘busy, business’ and *sōmer*, *sōmeres* ‘summer (sg., pl.)’. As far as I am aware, Anderson’s is a unique and uniquely compelling argument for disjunctive ordering: in this case, conjunctive application of the more general OSL before the more specific TSS leads to incorrectly lowered short vowels in the relevant contexts, as shown by the following derivations.

(40) OSL conjunctively precedes TSS (after Anderson 1969:141, 1974:104)

- | <u>UR</u>       | → | <u>OSL</u> | → | <u>TSS</u> | → | <u>SR</u>    |
|-----------------|---|------------|---|------------|---|--------------|
| a. / bǐsi+nes / | → | bēsines    | → | bēsines    | → | *[ bēsines ] |
| b. / sūmer+es / | → | sōmeres    | → | sōmeres    | → | *[ sōmeres ] |

Anderson’s proposal is for TSS to disjunctively block both subcases of OSL, but since these two rules are not abbreviable via any devices associated with disjunctive ordering (see again footnote 7), the options are either to allow disjunctive ordering to be stipulated *ad hoc* or to modify the conditions for disjunctive rule application. Anderson (1969:141-143, 1974:105), rather tentatively, pursues the latter route, pointing out that at least the structural descriptions of TSS and OSL can be abbreviated as  $\_ \text{ CV(C}_0\text{V)}$ .

[T]he rule corresponding to the longer of the two environments in the expansion of [  $\_ \text{ CV(C}_0\text{V)}$  ] takes precedence [...]. [W]hile the structural changes performed by the two rules are not identical, as they would have to be in order to allow the parentheses notation as presently defined to be appealed to, [...] both serve primarily to specify the value of the feature of length. [...] [T]he parentheses notation could be modified, so as not to require total identity of strings [...] but rather that (a) the structural changes of the two rules be ‘related’; and (b) the structural descriptions be collapsable by the present convention for the use of parentheses. (Anderson 1974:105)

Though Anderson does not define the sense in which the complex change specified by OSL is “primarily” about length as opposed to height, his point is nevertheless clear: whatever else they do, OSL and TSS demand opposite values of [±long] in the set of contexts in which both rules apply, and thus their structural changes are “related”.<sup>27</sup>

<sup>27</sup> It is worth contemplating what an OT analysis of the Middle English facts as presented by Anderson would have to involve; in particular, how lowering will also be blocked by shortening, given that these changes are not *prima facie* incompatible. Given an incompatible change  $\Delta_i$  that is blocked, a compatible change  $\Delta_c$  is also blocked iff the conditions for  $\Delta_c$ ’s application are only met when  $\Delta_i$  applies. Lowering must thus be dependent on lengthening in this way, as independently argued by Schane (1984).

#### 4.1. Enter the EC

Why go elsewhere to be cheated? Come here first!  
— *misleading used car advertisement*

The passage in Anderson’s work that comes closest to resembling Kiparsky’s (1973) statement of the EC, with which more linguists are familiar, is as follows:

Given two rules, if the domain of one is a proper subset of the set of forms to which the other can apply, we can refer to the former as the ‘more specific’ rule. We can then say that if two rules are adjacent, and specify changes that are related, then for any given form, if the more specific of the two is applicable, this precludes the application of the other. (Anderson 1974:107)

Kiparsky’s (1973:94) original statement of the EC is reproduced in (41) below; the main difference is that Kiparsky fines “related” changes as those that are “identical or incompatible”.<sup>28</sup> Both Anderson (1969:143, 1974:107) and Kiparsky (1973:94, fn. 3) note that the roots of the EC can be traced back to Pāṇini’s grammar;<sup>29</sup> Prince & Smolensky’s (1993/2004) PTC, of course, eponymously acknowledges this lineage.

(41) The Elsewhere Condition (Kiparsky 1973:94)

Two adjacent rules of the form

$A \rightarrow B / P \_ \_ Q$

$C \rightarrow D / R \_ \_ S$

are disjunctively ordered if and only if:

- a. the set of strings that fit  $PAQ$  is a subset of the set of strings that fit  $RCS$ , and
- b. the structural changes of the two rules are either identical or incompatible.

Rule pairs are identified as being subject to the EC through apparently simple comparisons of the rules’ structural descriptions and structural changes. The structural description inclusion requirement (41)a of the EC is met by any pair of rules that share a necessary condition that is sufficient for one rule but not for the other. As noted earlier, Nasalization and Oralization meet this requirement, as do Lengthening and Shortening in Modern English, as do OSL and TSS in Middle English. The (identical or) incompatible structural changes requirement (41)b also holds of these

---

<sup>28</sup> As Howard (1975:110) points out, Kiparsky does not define these terms and “thus their meanings must be inferred. On the basis of the examples cited, it seems reasonable to conclude that incompatible changes involve contrary specifications for the same feature, or assimilation and deletion.”

<sup>29</sup> Kiparsky also cites a “similar principle” in a prepublication version of Koutsoudas *et al.* (1974). As noted in footnote 5, Koutsoudas *et al.* invoke this principle (the PIPP) to predict the order between certain mutually bleeding rules. However, they note that mutually bleeding rules are “intrinsically disjunctive, since application of either rule yields a representation that fails to satisfy the structural description of the other. However, [...] there can be proper inclusion precedence between rules which do not destroy each other’s contexts, but which must nevertheless be prevented from applying conjunctively [...]. It thus seems appropriate to consider disjunctive application as part of the meaning of proper inclusion precedence” (p. 9, fn. 7). The version of the PIPP that would result would be in all essential respects identical to the EC.

three pairs of rules, since the two rules in each case make opposite, and hence incompatible, demands on some feature.

Consider how this works for the case of Lengthening and Shortening. The more specific structural description of Lengthening is met by words like *re(médi)⟨al⟩*, because the nonhead vowel of the branching foot is an /i/ in hiatus. So, Lengthening applies; the more general structural description of Shortening is also met here, but it is blocked from applying by the EC. (Based on the corresponding vowel in the underived word *(ré)mě⟨dy⟩*, the accented vowel in *re(médi)⟨al⟩* is assumed to be underlyingly short.)

(42) Lengthening blocks Shortening

$$\begin{array}{cccc} \underline{UR} & & \underline{Lengthening} & & \underline{Shortening} & & \underline{SR} \\ /re(médi)⟨al⟩/ & \rightarrow & |re(médi)⟨al⟩| & \rightarrow & blocked & \rightarrow & [re(médi)⟨al⟩] \end{array}$$

The structural description of Lengthening is not met by words like *(týpi)⟨cal⟩*, for two separate but equivalent reasons: the head vowel of the branching foot is [+high], and the nonhead vowel is not in hiatus. However, the structural description of Shortening is met by such words. (Based on the corresponding vowel in the underived word *(tý)pe*, the accented vowel in *(týpi)⟨cal⟩* is assumed to be underlyingly long.) Since Lengthening does not apply, it does not block Shortening from correctly applying.

(43) Shortening applies because Lengthening is not applicable

$$\begin{array}{cccc} \underline{UR} & & \underline{Lengthening} & & \underline{Shortening} & & \underline{SR} \\ / (týpi)⟨cal⟩ / & \rightarrow & n/a & \rightarrow & |(týpi)⟨cal⟩| & \rightarrow & [(týpi)⟨cal⟩] \end{array}$$

#### 4.2. Stipulation vs. explanation

Careful examination of the various component parts of the EC reveals that there is far more *ad hoc* stipulation than explanation involved in its formulation. In some cases, proponents of the spirit of the EC have discovered problems (empirical or otherwise) with one or more of these component parts and have modified the letter of the EC accordingly. As the EC has evolved through these modifications, its empirical coverage has arguably improved but its theoretical foundation is no more secure than it was in its original formulations in Anderson (1969, 1974) and Kiparsky (1973). Though these modifications were in most cases correctly made in response to perceived deficiencies in the EC, the fact remains that no part of the EC is derivable from independent theoretical considerations: each and every component of the EC could be defined otherwise, with the consequences being strictly and directly empirical. The question that proponents of the EC seem to have been asking in their investigations is: *What is the EC?* The question that I would argue should be asked is: *Why is the EC the way it is, and not some other way?*

As I demonstrate in §4.4 further below, this is what separates OT's PTC from the EC: because PTC is a theorem as opposed to an independently stipulated principle, its formulation could really be no other way unless the basic theory of constraint interaction in which it is couched is fundamentally changed. Not so in the case of the EC: in each of the following six subsections, I discuss a distinct stipulated component of the EC.

#### 4.2.1. Identical changes

The first component of the EC that we will consider is the proviso that the changes made by the two rules in question be “identical or incompatible”. Kiparsky (1982:8) drops the “identical or” portion of this requirement, stating instead that “[t]he result of applying [the more specific rule] *is distinct from* the result of applying [the more general rule]” (emphasis added). An explanation is provided in note 2 (Kiparsky 1982:84):

[A]s pointed out by Howard (1975), the case where the changes are identical was only necessary for stress rules. For example, the two rules collapsed in the schema  $V \rightarrow \acute{V} / \_ C_0 (V C_0) \#$  must apply disjunctively so that a stress is assigned to the final syllable only if there is no penult (i.e. in monosyllables). However, if we adopt metrical phonology, the two stress patterns [...] are distinct if we construe distinctness for metrical structure in the obvious way as incompatibility of labeling or bracketing. And in any case, the metrical version of the rule simply assigns a maximally binary foot to the right edge of a word, and so does not properly constitute a schema abbreviating two rules. We therefore need only specify distinctness of outputs [...].

Thus, the apparently disjunctive application of stress rules is either reduced to incompatibility or involves metrical assignment mechanisms for which ordering, disjunctive or otherwise, need not be specified.<sup>30</sup> Most if not all subsequent work on the EC appears to agree on the correctness of this modification, even in cases where the authors appear to explicitly adopt the formulation of the EC in Kiparsky (1973) as quoted in (41) above. For example, Halle (1995:27) reproduces the “identical or incompatible” requirement, but notes that “Kiparsky’s stress examples [...] have a solution that requires no reliance on the Elsewhere constraint [*sic*]”. This is either an oversight, or else Halle is implicitly pointing to the existence of cases in which rules that make identical changes other than stress must be disjunctively ordered by the EC. I know of no such examples.

Requiring that the structural changes of rules be only incompatible appears to be a generalization of the EC in that there is no longer the relevant disjunction among cases to which the EC is applicable (identical *or* incompatible structural changes). One could also interpret this as a less general formulation, since the EC now applies to fewer types of cases, though this interpretation is mitigated to the extent that the identical-change cases all are, and could only be, stress rules. But what this issue reveals is that the condition on structural changes is, from a strictly theoretical point of view, entirely arbitrary: why must the changes be incompatible, other than to cover the correct set of cases?

In OT, on the other hand, the restriction to incompatible changes is natural: incompatibility of the changes demanded by markedness constraints is what makes markedness constraints conflict with each other, and thus what requires them to be (directly) ranked with respect to each other. If two markedness constraints were sat-

---

<sup>30</sup> Paradoxically, Johnson (1972:113ff) proposes to dispense with the SPE parenthesis notation because “parentheses are not formally associated with the stress feature. Indeed, with their aid we can formulate many disjunctively applied rules that have nothing to do with stress. In this respect the principle of disjunctively [*sic*] ordering is grossly over-generalized, for it seems to be properly associated only with certain rules which introduce primary accent (whether of pitch or stress). Supposed cases of disjunctive ordering that do not fall under this rubric appear to be spurious.” See also footnote 37 below.



ified by identical changes to a form, they could both be satisfied without any need for ranking.

#### 4.2.2. Rule adjacency

Another stipulated component in both Anderson's and Kiparsky's original formulations of the EC is that two rules can be disjunctively ordered only if they are *already adjacent* in the ordering. As Prince (1997a:3, note 4) notes, the adjacency requirement appears to be nothing more than a hold-over from SPE, since this prerequisite was "obviously necessary for reduction via parentheses". That the adjacency requirement may be unnecessary is already acknowledged by Kiparsky (1973:104), who comments that the EC "might be applicable not just to rules which are adjacent in the ordering, but also to rules which are separated by other rules." Kiparsky goes on to discuss a potential case bearing on this issue, Halle's (1971) account of accent assignment in Slavic; because the relevant rules make identical changes of the stress type, the right analysis probably involves metrical structure assignment rather than the EC (recall the discussion in §4.2.1 just above).

In his analysis of the interaction between (change-wise incompatible) harmony and lowering in Brazilian Portuguese, Harris (1974) argues that harmony is crucially separated from lowering by truncation and stress, and that yet harmony disjunctively blocks lowering. Kiparsky (1982:8), without specific comment, drops the adjacency requirement and instead requires more loosely that the rules be in the same component.<sup>31</sup>

(44) The Elsewhere Condition (Kiparsky 1982:8)

Rules A, B in the same component apply disjunctively to a form  $\Phi$  if and only if

- (i) The structural description of A (the special rule) properly includes the structural description of B (the general rule).
- (ii) The result of applying A to  $\Phi$  is distinct from the result of applying B to  $\Phi$ .

In that case, A is applied first, and if it takes effect, then B is not applied.

Again, we are confronted with an issue of theoretical arbitrariness: no adjacency or componential affiliation condition is justifiable outside of its empirical consequences. In OT, of course, conflict between two markedness constraints holds regardless of the adjacency of the constraints in the constraint hierarchy.

#### 4.2.3. Rule precedence

Once the rule adjacency requirement is abandoned, the question arises whether the EC imposes (or presupposes) a particular linear order between the rules: the more

---

<sup>31</sup> As already noted in §2.2, this entails that Lengthening and Shortening in English, as discussed by Kiparsky (1982:44), are not disjunctively ordered by the EC but rather are conjunctively ordered due to the inherently conjunctive order between their separate components: cyclic for Shortening, postcyclic for Lengthening. Kenstowicz (1994a:217ff) cites Kiparsky's (1982) version of the EC and yet uses Lengthening and Shortening as an example of how the EC works. Halle's (1995) statement of the EC is essentially Kiparsky's (1973) version but without the adjacency requirement, but there is also no discussion of how Lengthening and Shortening are related to each other in terms of adjacency (or componential affiliation).

specific rule might be required to precede (and if it applies, to block) the more general rule, or vice-versa — or neither order between the rules may be required to hold.

Kiparsky's (1982) version of the EC in (44) above stipulates the first of these three options: "A is applied first, and if it takes effect, then B is not applied." As far as I am aware, Kiparsky presents no argument for this position over the logical alternatives. As noted by Prince (1997a:3, note 4), Goldsmith's (1981, 1984ab) analysis of Tongan involves blocking of the more general rule which must otherwise precede the more specific rule; in response, Halle & Vergnaud (1982:81) counter that Goldsmith's analysis "is quite unprecedented in the literature" and continue their disagreement with it as follows:

[T]he disjunctivity required is not the traditional one but, to quote Goldsmith, rather "one according to which the more GENERAL rule can be precluded from applying not only by the ACTUAL application of the more specific rule, but [by] the PRESENCE of the more specific rule — that is, by its potential application later in the derivation" [...]. These are far-reaching modifications in the algorithm for rule application and they are not to be welcomed since they render the rules extremely powerful, for now whether or not a given rule applies to a given string no longer depends purely on the form of the string, but may also depend on what other rules might apply to the input string.

While it is indeed the case that Goldsmith's analysis requires that the earlier-ordered more general rule somehow 'know' whether the later-ordered more specific rule might apply to the string, whether this analysis can be considered "unprecedented in the literature" depends on how broadly we construe the analysis and the literature. It may be true that Goldsmith's suggestion was unprecedented in the literature *on the EC* (and disjunctive ordering more generally) up to that point, but a wealth of work in the early 1970s (Kisseberth 1970, Hill 1970, among many others), Halle & Vergnaud's possible opinions of it notwithstanding, had established a clear precedent for at least the *possibility* that relations among rules might have access to information that is not strictly in their input strings. Moreover, as I document in §4.2.6, Halle himself in later work suggests a modification to the EC with similar consequences for "the algorithm for rule application".

The analysis of elsewhere interactions in OT does require a particular ordering relationship between the more specific and more general constraints; namely, that the more specific dominate the more general, in order for the more specific constraint to have any effect at all (see §4.4). This means that the analog of Goldsmith's analysis in OT, one in which the more specific constraint blocks the more general constraint even though the latter dominates the former, would indeed be impossible as a matter of definition.

#### 4.2.4. External context vs. whole structural analysis inclusion

There is another potential modification of the EC noted by Kiparsky (1973:104):

[A] subset relationship in the external context—namely,  $P\_Q, R\_S$  in [(41)]—rather than the whole structural analysis— $PAQ, RCS$ —might suffice to establish disjunctivity.

Adopting Kiparsky's terms, I refer to this modification as *external context inclusion*, as opposed to the original *whole structural analysis inclusion*. Kiparsky briefly and speculatively discusses an example in Sanskrit in which only external context inclusion, not whole structural analysis inclusion, holds of the rules in question. The

two rules are stated in (45). (Note that [ $\alpha$ place] in (45)a refers to subcoronal places of articulation; i.e., those minor place distinctions usually covered by [ $\pm$ anterior] and [ $\pm$ distributed].)

(45) Sanskrit rules (Kiparsky 1973:104-105)

- a. Place Assimilation (PA) b. Fricative Aspiration (FA)
- $$[+\text{coronal}] \rightarrow [\alpha\text{place}] / \_ \# \left[ \begin{array}{l} \alpha\text{place} \\ +\text{coronal} \\ -\text{continuant} \end{array} \right] \quad \left[ \begin{array}{l} +\text{coronal} \\ +\text{continuant} \end{array} \right] \rightarrow h / \_ \#$$

Both rules apply to word-final coronals, but PA in addition requires a following (word-initial) coronal stop to for the word-final coronal assimilate to; thus, the external context of PA includes that of FA. However, FA applies more specifically to coronal fricatives, meaning that the whole structural analysis of PA does not include that of FA. Nevertheless, Kiparsky claims, the rules must apply disjunctively: if the more specific external context of PA holds, then PA applies, blocking FA (e.g.,  $s\#c \rightarrow s\#c$ ,  $*h\#c$ ).<sup>32</sup> The apparent problem in this case is that PA applies to both coronal fricatives *and* coronal stops; whole structural analysis inclusion could be maintained if PA were simply split into those two halves, with the half applying to fricatives disjunctively blocking FA.<sup>33</sup>

As discussed at length in §4.3 below, an appeal to simplicity (SPE's evaluation procedure) cannot technically be made to rule out either of these proposed reanalyses. If whole structural analysis inclusion is correct, then Kiparsky's and Harris's rules apply conjunctively and the relevant rules in the proposed reanalyses apply disjunctively; each analysis thus describes a different set of facts, about which the evaluation procedure has nothing to say. It is significant, then, that the desire for disjunctive application in both of these cases forces either (i) the splitting of the more specific rule into two or (ii) the dubious relaxation of the EC's inclusion requirement, neither of which is very appealing. There is, however, a type of case for which reanalysis along these lines is impossible and thus in which relaxation of the EC's inclusion requirement appears to be necessary.

A specific example is Halle & Idsardi's (1997) analysis of *r*-deletion and *r*-insertion in Eastern Massachusetts English, which generalizes to all cases of complementary deletion and insertion rules.<sup>34</sup> Because deletion involves a change  $x \rightarrow \emptyset$  and insertion involves a change  $\emptyset \rightarrow x$ , it is impossible for the whole structural analysis of both rules to be in a subset relationship. For example, in Eastern Massachusetts English, *r*-insertion is the more specific rule  $\emptyset \rightarrow r / P \_ Q$  and *r*-deletion is

<sup>32</sup> Kiparsky notes that coronal fricatives also optionally assimilate to coronal fricatives and consonants at other places of articulation ( $s\#\xi \rightarrow s\#\xi$  or  $h\#\xi$ ), in which case FA must also be blocked.

<sup>33</sup> Harris (1974:77ff) also uses external context inclusion to establish disjunctivity between harmony and lowering in Brazilian Portuguese. The same reanalysis is available here: the 'more specific' harmony rule could be split in two, one half applying to just those vowels to which lowering applies.

<sup>34</sup> The point made here is also made by McCarthy (1999b:7). Halle & Idsardi's (1997) analysis is discussed in more detail, with specific reference to a separate issue in their redefinition of the EC, in §4.2.6 below. (Note the Halle & Idsardi do not themselves explicitly note the inclusion requirement issue.)

the more general rule  $r \rightarrow \emptyset / R \_ S$ . The whole structural analysis of  $r$ -insertion is thus  $P\emptyset Q$ , and the whole structural analysis of  $r$ -deletion is  $RrS$ . Thus, no subset relationship can be established between these whole structural analyses, no matter what  $P$ ,  $Q$ ,  $R$ , and  $S$  are.

Thus if we want cases of complementary deletion and insertion rules to be disjunctively ordered by the EC, external context inclusion seems a welcome modification. There are two problematic issues that this modification raises, however. The first is the fact that the  $A \rightarrow B / P \_ Q$  rule format is only intended to be a notational shorthand for the rewrite transformation  $PAQ \rightarrow PBQ$ ; slash and dash ('/', '\_') are "auxiliary expressions" in the SPE formalism (Chomsky & Halle 1968:393). Second, and more seriously, with external context inclusion it is unclear what relationship, if any, must hold between the *foci* of the two rules ( $A$  and  $C$  in Kiparsky's (1973) definition). The only clause of the EC that might now make any reference to  $A$  and  $C$  is the one requiring that the structural changes of the rules ( $A \rightarrow B$ ,  $C \rightarrow D$ ) be incompatible, or more specifically, that the application of one of the rules to a form  $\Phi$  be distinct from application of the other to  $\Phi$ . But in the case of complementary insertion and deletion rules, there is no form  $\Phi$  to which both rules apply! A form with the relevant element in the right context can only undergo deletion, and a form without it in the right context can only undergo insertion. Reformalization of the correct sense of 'incompatible/distinct changes' is clearly needed.

Even if it is possible to address these issues in a principled manner, the choice between whole structural analysis and external context inclusion is still a theoretically arbitrary one; it is yet another stipulation that remains to be explained. An OT analysis of complementary insertion and deletion processes involves no such arbitrariness: because markedness constraints evaluate output strings, any markedness constraint  $M_1$  requiring insertion of  $x$  necessarily conflicts with a markedness constraint  $M_2$  requiring deletion of  $x$  if and when their contexts of application overlap. The only issue for an OT analysis is ensuring that the same segment is inserted and deleted; see Baković (1999) for discussion of this point with specific reference to the Eastern Massachusetts English  $r$  case.

#### 4.2.5. Substring vs. whole-form disjunction

There is a subtle ambiguity in what it means for a rule to apply to a form that has important consequences for the interpretation of the EC, assuming now Kiparsky's (1982) definition in (44) above. Suppose that a given form  $\Phi$  has multiple substrings (say two,  $s_1$  and  $s_2$ ) that meet the structural description of the general rule B, and that some but not all of those substrings also meet the structural description of the special rule A (say, only  $s_1$ ). Assuming that the structural changes made by A and B are incompatible in the relevant sense means that the result of applying A to  $\Phi$  is distinct from the result of applying B to  $\Phi$ ; applying B affects both  $s_1$  and  $s_2$ , while applying A affects only  $s_1$ .<sup>35</sup>

---

<sup>35</sup> Of course, A and B affect  $s_1$  differently. But technically, under Kiparsky's (1982) 'distinctness of application to  $\Phi$ ' definition, A and B *could* affect  $s_1$  identically and yet still be disjunctively ordered by the EC for just those forms in which there is an  $s_2$  that only B affects! This is clearly the wrong result.

How does disjunction work in this case? Clearly, B must be blocked from applying to  $s_1$ , or else B would simply undo the effect of A's application; call this *substring disjunction*. But what about  $s_2$ ? Conceivably, application of A to  $s_1$  could block application of B such that B applies to neither  $s_1$  nor  $s_2$ . Call this *whole-form disjunction*.

For concreteness, recall the disjunctive application of the Nasalization (= A) and Oralization (= B) rules in §2.1.5. Suppose we have a set of underlying representations of the form /NVOV/ and /OVOV/, with different specifications for [ $\pm$ nas] on the vowels. Nasalization is applicable to the substring NV and Oralization is applicable to both NV and OV. The consequences of blocking by substring disjunction are shown in (46). Blocking of Oralization on the substring  $NV_N$  to which Nasalization has applied is indicated with underlining; Oralization is free to apply to the substring OV, rendering  $OV_o$ .

(46) Blocking by substring disjunction

	<u>UR</u>		<u>Nasalization</u>		<u>Oralization</u>		<u>SR</u>
a.	/NV <sub>N</sub> OV <sub>N</sub> /	→	NV <sub>N</sub> OV <sub>N</sub>	→	<u>NV<sub>N</sub></u> OV <sub>o</sub>	→	[NV <sub>N</sub> OV <sub>o</sub> ]
b.	/NV <sub>N</sub> OV <sub>o</sub> /	→	NV <sub>N</sub> OV <sub>o</sub>	→	<u>NV<sub>N</sub></u> OV <sub>o</sub>	→	[NV <sub>N</sub> OV <sub>o</sub> ]
c.	/NV <sub>o</sub> OV <sub>N</sub> /	→	NV <sub>o</sub> OV <sub>N</sub>	→	<u>NV<sub>o</sub></u> OV <sub>o</sub>	→	[NV <sub>o</sub> OV <sub>o</sub> ]
d.	/NV <sub>o</sub> OV <sub>o</sub> /	→	NV <sub>o</sub> OV <sub>o</sub>	→	<u>NV<sub>o</sub></u> OV <sub>o</sub>	→	[NV <sub>o</sub> OV <sub>o</sub> ]
e.	/OV <sub>N</sub> OV <sub>o</sub> /	→	n/a	→	OV <sub>o</sub> OV <sub>o</sub>	→	[OV <sub>o</sub> OV <sub>o</sub> ]
f.	/OV <sub>o</sub> OV <sub>N</sub> /	→	n/a	→	OV <sub>o</sub> OV <sub>o</sub>	→	[OV <sub>o</sub> OV <sub>o</sub> ]

The surface distributions of nasal and oral vowels in (46) are exactly what we expect: nasal vowels after nasals and oral vowels elsewhere. Now consider the consequences of blocking by whole-form disjunction, as shown in (47). Here, Oralization is blocked from applying to the entire form whenever Nasalization is applicable anywhere in the form, as in (47)a-d, but Oralization is free to apply to OV substrings in just those forms to which Nasalization is nowhere applicable, rendering  $OV_o$  as shown in (47)e-f.

(47) Blocking by whole-form disjunction

	<u>UR</u>		<u>Nasalization</u>		<u>Oralization</u>		<u>SR</u>
a.	/NV <sub>N</sub> OV <sub>N</sub> /	→	NV <sub>N</sub> OV <sub>N</sub>	→	<i>blocked</i>	→	[NV <sub>N</sub> OV <sub>N</sub> ]
b.	/NV <sub>N</sub> OV <sub>o</sub> /	→	NV <sub>N</sub> OV <sub>o</sub>	→	<i>blocked</i>	→	[NV <sub>N</sub> OV <sub>o</sub> ]
c.	/NV <sub>o</sub> OV <sub>N</sub> /	→	NV <sub>o</sub> OV <sub>N</sub>	→	<i>blocked</i>	→	[NV <sub>o</sub> OV <sub>N</sub> ]
d.	/NV <sub>o</sub> OV <sub>o</sub> /	→	NV <sub>o</sub> OV <sub>o</sub>	→	<i>blocked</i>	→	[NV <sub>o</sub> OV <sub>o</sub> ]
e.	/OV <sub>N</sub> OV <sub>o</sub> /	→	n/a	→	OV <sub>o</sub> OV <sub>o</sub>	→	[OV <sub>o</sub> OV <sub>o</sub> ]
f.	/OV <sub>o</sub> OV <sub>N</sub> /	→	n/a	→	OV <sub>o</sub> OV <sub>o</sub>	→	[OV <sub>o</sub> OV <sub>o</sub> ]

Both intuitively and empirically, whole-form disjunction produces a bizarre result. The distribution of nasal and oral vowels described is: vowels are nasal after nasals and oral elsewhere, except a nasal/oral distinction surfaces in otherwise oral contexts just in case there is also a nasal context somewhere else in the same form.<sup>36</sup>

<sup>36</sup> Of course, cases of unbounded complementary distribution like the intended one here have the luxury of alternative analyses in terms of prior underlying specification (1) or under-specification + feature-filling rule application (2). The same point could be made with a case of bounded complementary distribution, which does not have either of these analytical re-

As a historical note: Anderson (1974:108-109), citing Bever (1967), Howard (1972), and Johnson (1972), briefly addresses — but does not resolve — this very issue.<sup>37</sup>

[W]hile [the theory of disjunctive ordering] specifies the conditions under which the application of one rule precludes the application of some other, it was only meant to cover cases in which the domains of the two rules overlapped to some extent. [...] [W]e leave the issue unresolved at this point, except to note that disjunctive ordering is presumed not to obtain if the portion of the form analyzed by the [structural description] of [the more specific rule] is totally disjoint from the portion analyzed by the [structural description] of [the more general rule].

The issue, yet again, is that either form of disjunction is a logical possibility and the empirically correct one remains to be explained rather than stipulated. Notably, only substring disjunction is possible in OT. Suppose that a form has two substrings  $s_1$  and  $s_2$  to which the more general constraint  $\mathbb{G}$  is applicable, but that the more specific constraint  $\mathbb{S}$  is only applicable to  $s_1$ . If  $\mathbb{S}$  dominates  $\mathbb{G}$ , then  $\mathbb{G}$  will be forced to be violated only by  $s_1$ , the substring to which both constraints are applicable; since  $\mathbb{S}$  is irrelevant to the other substring,  $\mathbb{G}$  will emerge victorious there. This is due to the *minimal violation* property of OT (Prince 1993): a constraint is not violated unless it is specifically forced to be violated due to conflict with a higher-ranking constraint. Any attempt to implement whole form disjunction in OT would have extensive undesirable consequences; for example, deletion of a consonant due to NOCODA  $\gg$  MAX would parasitically license deletion of other segments penalized by markedness constraints ranked below MAX.

#### 4.2.6. Applicational vs. situational disjunction

Citing the paper that became Halle & Idsardi (1997), Halle (1995:27) advances yet another modification to the definition of disjunctive ordering by the EC (emphasis added):

[T]he less restrictive rule may not apply to a **string that has the form of the more restrictive rule**. [This definition] is somewhat more general than the one given in Kiparsky (1973), where “disjunctivity” was limited to **strings to which the more “restrictive” rule had applied**. [...]

I refer to the definition of disjunctive ordering attributed to Kiparsky (1973) here as *applicational disjunction*, and to Halle’s modified definition as *situational disjunction*. The empirical base of Halle (1995) is the interaction of Lengthening and Shortening in English, and so the discussion of the differences between these two

---

courses (as discussed in §2.2). I chose the unbounded complementary distribution case for the purposes of demonstration due to its relative straightforwardness.

<sup>37</sup> In the most extensive of these discussions, Howard (1972:78ff) distinguishes two types of substring disjunction: *environmental disjunction* —equivalent to Bever’s (1967:110ff) *segmental disjunction* — and *parenthetical disjunction*. (Whole-form disjunction is referred to as *word-level disjunction*.)

Johnson (1972) is a minor revision of his 1970 doctoral dissertation; Anderson cites the dissertation. Note that Johnson (1972:114ff) simply presupposes whole-form blocking in his argument against the use of the parenthesis notation (see footnote 30 above), and does not appear to consider substring blocking.

definitions of disjunction begins there before moving on to the empirical base of Halle & Idsardi (1997).<sup>38</sup>

What Halle (1995) intends to achieve with situational disjunction is for Shortening not to apply to any long vowel in the Lengthening environment, regardless of whether that vowel is long due to an actual (that is, nonvacuous) application of Lengthening or due to its being underlyingly long. As evidence for the correctness of situational disjunction as opposed to applicational disjunction, Halle (1995:28) notes that the heads of branching feet in words like *Shake*(*spēāri*)⟨*an*⟩ and ⟨*jōvi*⟩⟨*al*⟩ are not subject to Shortening “even though Lengthening is not responsible for the long vowel in their stems.”

Halle either assumes that Lengthening must be stated so as to apply to short vowels only, or else that nonvacuous application does not ‘count’ as application; it is impossible to know which, as Halle does not discuss the issue (nor, in fact, does he formally state either of the two rules). Note, however, that in order for Lengthening and Shortening to satisfy the whole structural analysis inclusion requirement of the EC (recall §4.2.4 above), they must both be stated in such a way that their foci are not themselves distinct. Both rules must thus apply to any vowel, not just short ones for Lengthening or long ones for Shortening; therefore, Lengthening must apply, albeit vacuously, to words like *Shake*(*spēāri*)⟨*an*⟩ and ⟨*jōvi*⟩⟨*al*⟩. Even if the EC were to more weakly require external context inclusion, there does not appear to be anything to prevent the rules from being stated so that they both apply to all vowels; indeed, the SPE evaluation procedure would appear to require this generalization in their formal statement. So, unless there is independent reason why vacuous application should not ‘count’ as application for the EC to enforce disjunction — and Halle provides no such reason — applicational and situational disjunction cannot strictly be distinguished by this case.

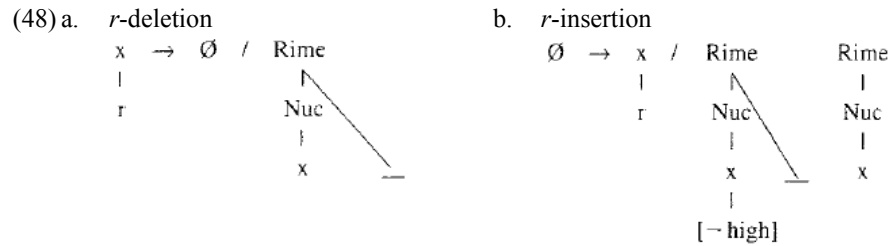
Halle & Idsardi’s (1997) case of *r*-insertion and *r*-deletion in Eastern Massachusetts English (henceforth EME) does distinguish the two types of disjunction, and indeed requires situational disjunction. In EME, *r* is deleted pre-consonantly and word-finally (in the rime or coda) but *r* is also inserted pre-vocally after nonhigh vowels across a word boundary. Of interest here is the fact that, restricting our attention to word-final contexts, *r* is inserted after a nonhigh vowel and before a word-initial vowel; elsewhere, *r* is deleted.<sup>39</sup> Halle & Idsardi’s (1997:343-344) rules are reproduced in (48).<sup>40</sup>

---

<sup>38</sup> This same issue in relation to Halle & Idsardi (1997) is discussed by McCarthy (1999b:7-8).

<sup>39</sup> Note that this is a case of bounded complementary distribution, rendering SPE analyses like (1) and (2) inadequate (*pace* Pullum 1976:90-91). There are additional factors that are not germane here; see Kahn (1976), McCarthy (1991, 1993, 1999b), Harris (1994), Halle & Idsardi (1997), and references therein.

<sup>40</sup> Note that *r*-insertion in (48)b is carefully stated as insertion of a *coda* consonant — a typologically curious condition on a hiatus-breaking insertion rule, but crucially necessary in order for the external context of this rule to properly include the external context of *r*-deletion (48)a and to call the EC to action (recall §4.2.4 above, and see also footnote 41 and §4.3 below). Halle & Idsardi (1997:334), citing McCarthy (1993:179), point to the phonetic similarity of inserted *r* to true coda *r* and the opposition of both to true onset *r*; in McCarthy (1993),



Based on a comment by François Dell, Halle & Idsardi (1997:345) note that application of the more general *r*-deletion rule (48)a must be blocked not only if the more specific *r*-insertion rule (48)b has applied, but also if the output of *r*-insertion just so happens to be present because it surfaced unchanged from the underlying representation (or indeed for any other reason, though this is perhaps beside the point here). Situational disjunction is necessary here to account for examples of word-final *rs* in the context of *r*-insertion that are not deleted but that were certainly not inserted; i.e., the ‘linking’ *r* in e.g. *Homer**r* arrived, as opposed to the ‘intrusive’ (= inserted) *r* in e.g. *Wandar**r* arrived. Since *r*-insertion cannot be stated so as to have applied, even vacuously, to forms with linking *r*, its application alone cannot suffice to block *r*-deletion; the independent existence of a form that could have resulted from a nonvacuous application of *r*-insertion must also be able to block application of *r*-deletion.<sup>41</sup>

Just as with the external context vs. whole structural analysis inclusion (§4.2.4), the distinction between situational and applicational disjunction appears to be necessary only in cases of complementary deletion and insertion rules. In the case of situational vs. applicational disjunction, this is because deletion and insertion rules do not apply vacuously. One must ask whether things could be otherwise. If it happened that disjunction were applicational and not just situational, then we would expect complementary deletion and insertion rules to behave in a very different manner from other complementary rules. In the case at hand, the result would be that intrusive *r* would exist but linking *r* would not: the *r* in *Homer**r* arrived would be deleted because it wasn’t inserted, while the *r* in *Wandar**r* arrived would not be deleted because it couldn’t be, having been inserted. This is perfectly plausible alternative state of affairs; the fact that complementary deletion and insertion rules behave just like other complementary rules in actuality is simply serendipitous, a complete and very curious accident according to the EC ‘theory’.

following Kahn (1976), the inserted *r* is ambisyllabic (a sort of compromise response to both ONSET and FINAL-C), which can also account for the relevant phonetic facts.

<sup>41</sup> There is another not-so-small matter concerning Halle & Idsardi’s (1997) formulation of *r*-insertion in (48)b that may lead to some confusion here. As stated, the rule appears to only apply if there is a coda already, which means it should apply in the (vacuous) linking *r* case and not in the nonvacuous case of hiatus where it is actually necessary. (The rule might also be interpreted so as to apply to replace any other coda consonant in the relevant context, but I digress.) Creation of the syllable coda should in fact be a part of the structural change of the rule, or perhaps the consequence of some independent rule — but then we are faced even more obviously with the issue noted in footnote 40 above, that Halle & Idsardi state *r*-insertion very carefully so that it disjunctively blocks *r*-deletion in the correct way.



Halle & Idsardi (1997:346-347) respond to a reviewer who asks whether situational disjunction “does not amount to a trans-derivational constraint” (or global rule):

The reinterpretation of the Elsewhere Condition is not a trans-derivational constraint because no access to previous or subsequent representations in the derivation (or other derivations) is necessary. Rather, the application of a rule to the current representation is blocked in one particular circumstance—if the form is compatible with the structural description or change of a more specific rule. Thus, what is required is a limited access to the formal encodings of other rules [...].

My view is that Halle & Idsardi (1997) are splitting hairs pretty finely here, especially since they do not define how situational disjunction must work. Here is an attempt at a semi-formal definition. Situational disjunction requires that the EC perform a special assessment of the form  $\mathcal{F}$  in the derivation that serves as input to the more general rule GR.<sup>42</sup> If  $\mathcal{F}$  corresponds to a nonvacuous output of the more specific rule SR, then GR is blocked from applying to  $\mathcal{F}$ ; if it doesn't, then GR applies to  $\mathcal{F}$ . In order to perform this assessment, the EC must somehow match  $\mathcal{F}$  with SR's structural change and external context ( $PBQ$ , given  $SR = A \rightarrow B / P \_ \_ Q$ ), blocking GR even though  $\mathcal{F}$  also matches GR's whole structural analysis ( $RCS$ , given  $GR = C \rightarrow D / R \_ \_ S$ ). This assessment mechanism is certainly quite powerful and global even if limited, something that Halle himself has admonished in other work (e.g., Halle & Vergnaud 1982; recall §4.2.3). Regardless, the issue here is the theme running through this section: the choice between situational and applicational disjunction is theoretically arbitrary.

In OT, markedness constraints are sensitive to output forms regardless of their input specification, so whether the process of lengthening (or insertion) ‘applies’ or not has no bearing on the ‘application’ of shortening (or deletion). Satisfaction of the higher-ranked, more specific of two conflicting markedness constraints simply has ranked priority over the lower-ranked more general of the two, causing the latter to be violated in just those cases where violation of the former is at stake. For example, violation of the higher-ranked, more specific constraint  $*[-high]_s / (\_ Ci)V$  is at stake in the assessment of words like *Shake*(*spēāri*)(*an*) and (*jōvi*)(*al*), regardless of the provenance of the long vowels, forcing violation of the lower-ranked more general constraint  $*V_L / (\_ \sigma)$ .

### 4.3. Rule formulation, the evaluation procedure, and derivation length

In the preceding subsection I identified six distinct stipulated components of the EC that, in Prince & Smolensky's (1993/2004) words, “could very well be otherwise.” The point has been to show that the EC is at best an approximation of a synthesis of the relevant facts and to explain how OT's fundamental assumptions entail the correct facts from the outset, without the need to stipulate anything more. The main point of this paper bears repeating here: ‘elsewhere’ relationships between processes fall out directly from OT's basic architecture, following from it as a theorem; the EC is a separate set of stipulations, each bearing no necessary relation to the SPE theory in which it is couched.

---

<sup>42</sup> For concreteness and simplicity, I assume here that substring disjunction is correct and that  $\mathcal{F}$  contains exactly one substring to which the more general rule applies (recall §4.2.5).

This is all the more significant when it is recognized that two rules that are disjunctively ordered by the EC can, at least in most cases, be easily rewritten so that they no longer meet the requirements of the EC. For instance, as noted in passing in §4.2.4, if Lengthening were stated so as to apply to short vowels only, or if Shortening were stated so as to apply to long vowels only, then whole structural analysis inclusion wouldn't hold.<sup>43</sup> Similarly, if *r*-insertion were stated so as to make the inserted *r* ambisyllabic (McCarthy 1993; see footnote 40), then neither external context nor whole structural analysis inclusion would hold. In both cases, the two rules would be predicted to apply conjunctively rather than disjunctively. This leaves open the problems associated with Duke of York derivations and calls into question the utility of the EC itself.

It is important to recognize that SPE's evaluation procedure (Chomsky & Halle 1968:330-335, 392-393) cannot be invoked to somehow disallow rules from escaping the EC. The evaluation procedure is a method for selecting among different grammars "which are all compatible with whatever data are available from a certain language" (Chomsky & Halle 1968:330).<sup>44</sup> Since a pair of rules will apply disjunctively if and only if they are written to meet the EC requirements and conjunctively otherwise, two grammars with the rules written in the two different ways under discussion describe different sets of data and are thus incommensurable in terms of the evaluation procedure.

The same point applies to notations that abbreviate disjunctively ordered sets of rules in general. Consider Chomsky & Halle (1968:71, note 16):

To say that the rules may be given in a simpler form implies that they *must* be given in that form. More precisely, the notations that we use define a certain valuation measure for grammars; the value of a grammar is determined by the number of symbols that appear in it when notations are used in the optimal fashion. Rules are ordered by conventions associated with the parenthesis (or other) notation when the use of this notation is in fact optimal in the case in question.

Any set of rules that is abbreviable via parentheses or angled brackets will apply disjunctively if and only if they are so abbreviated and conjunctively if and only if they are not, the result being (at least) two grammars that describe different sets of data, and which are thus incommensurable in terms of the evaluation procedure. Abbreviability of rules via parentheses or angled brackets must somehow be assessed independently of the evaluation procedure, perhaps as suggested by Chomsky & Halle (1968:63):

---

<sup>43</sup> Thanks to Roger Schwarzschild and Hubert Truckenbrodt for helping me to see this point, and to Alan Prince for discussion of it; see also Prince (1997a:3, note 5). Thanks also to John McCarthy for discussion of some of the broader implications, though he may not necessarily agree with my conclusions.

<sup>44</sup> Halle also reiterates this point in subsequent work: "Appeal to the simplicity metric [= the SPE evaluation procedure —EB] is irrelevant where alternative descriptions either cover different sets of facts or where they employ different theoretical devices" (Halle 1978:127/1979:331). Note also the historical view of the SPE stance on simplicity reported in Goldsmith & Laks (to appear, p. 7): "In case of multiple accounts of the same data [...] formal simplicity was to be used to choose among the accounts".

[I]t would be quite interesting to determine whether there is a general principle governing [the organization of the grammar in terms of disjunctive and conjunctive ordering]. A natural principle that suggests itself at once is this: *abbreviatory notations must be selected in such a way as to maximize disjunctive ordering*. [...] The principle seems to us a natural one in that maximization of disjunctive ordering will, in general, minimize the length of derivations in the grammar.

The passage from Chomsky & Halle (1968:63) quoted in §2.2.2 is a continuation of this one, and so the comments made there concerning the length of derivations and implementing the competence model in performance apply here.<sup>45</sup> However, as also noted in that section, there does indeed seem to be something to the idea that derivation length should be minimized, at least to the extent that (feeding) DY derivations will thereby be avoided (McCarthy 1999a, 2003) or, more generally, that rules that could undo each other's effects will thereby never apply to same element (Norton 2003).

As Norton (2003:136-137) notes, the EC is no real solution to this desideratum given that the one core element of the EC, the inclusion relationship between strings to which the rules apply, leaves out all other cases of mutually-undoing rules that apply in overlapping environments. This was also noted by Pullum (1976:94), who writes:

If it is to be required that no description contain an instance of the Duke of York gambit for any form generated, this will have to be accomplished by an *ad hoc* theoretical principle in the purely nonpejorative sense that the constraint will be set up for this special purpose. Independent constraints on the form of phonological descriptions [such as the EC —EB] will in some cases be in agreement with an anti-Duke of York principle and sometimes not.

The fact that the EC can be easily circumvented by rule rewriting thus highlights the inadequacy of the EC as an 'anti-DY principle' purportedly serving to shorten derivations. OT, which eschews derivations, eschews DY derivations *a fortiori*.

#### 4.4. On a recent invocation of the EC

In a recent paper, Piggott (2003) proposes an analysis of the typology of nasal harmony, which he argues requires massive revisions to OT; Piggott names the product of his revisions the "Principles & Parameters" (or "P&P") framework. The most sweeping of these revisions are (i) that not all constraints are present in every language, and (ii) that the ranking of the constraints present in a given language is determined by a set of metaconditions (Piggott 2003:406). Piggott (2003:401ff) names the only ranking metacondition he proposes in the paper "Pāṇini's theorem on constraint ranking", reproduced in (49).<sup>46</sup>

---

<sup>45</sup> Compare the following historical comment on SPE theorizing by Goldsmith & Laks (to appear, p. 7): "The number of rules whose ordering remains the same across all [total orderings consistent with each pairwise ordering empirically established] is the depth of the rule ordering, and it was a desideratum of the [SPE] framework to determine a large depth for a given language." This desideratum seems to be in direct conflict with the idea of minimizing the length of derivations; I have no idea if/how it is reconciled.

<sup>46</sup> On the relative inappropriateness of the Beckman citation, see §5.4 below.

## (49) Pāṇini's theorem on constraint ranking (PTCR; Piggott 2003:401-402)

Given a pair of constraints A and B, where B has scope over the set of elements X and A has scope over the set of elements Y where Y is a proper subset of X, A (the special case) takes precedence over B (the general case), if the demands of A and B conflict in some domain D (i.e. special  $\gg$  general).

Somewhat mysteriously, Piggott fails to mention Prince & Smolensky's theorem of the same name, and only cites Kiparsky (1973), Scobbie (1991, 1993), and Beckman (1997) as "precedence for the proposal" (Piggott 2003:401, fn. 14). But this is strangely appropriate: Piggott's "PTCR", which imposes a ranking on constraints, bears little resemblance to Prince & Smolensky's PTC, which does not. Given this important distinction, PTCR would be more appropriately named after the EC, especially given that it is not, in fact, stated as a theorem — *pace* Piggott (2003:401, fn. 14), who writes:<sup>47</sup>

This theorem expresses a logical necessity that is imposed on the P&P framework and is probably not a good candidate for a uniquely linguistic metacondition."

Although Piggott's intent with PTCR is somewhat clear, it is hard to make sense of such undefined terms as "has scope over", "elements", "demands of", and so on. In any event, what is crystal clear is that PTCR is analogous to the EC, not to PTC, and that it is thus subject to the same critique of the EC offered in the foregoing subsections.

### 5. Pāṇini's Theorem on Constraint-ranking

As noted several times already, PTC is a theorem as opposed to a principle; the proof of the theorem shows that it is a necessary consequence following from how constraints interact in OT. The theorem crucially involves the concept of *constraint activity*, and so it is useful to begin our discussion of PTC with some clarification of this critical concept.

Intuitively, a constraint  $\mathbb{C}$  is active in the selection of an optimal candidate from a given input if  $\mathbb{C}$  crucially rules out some competing candidate(s). Constraint activity is defined relative to two interrelated variables: an input and a constraint hierarchy.

Consider first the role of the input. According to the standard OT architecture of a grammar, the generator function Gen takes a given input  $i$  and returns a candidate set  $\text{Gen}(i)$  from which the optimal candidate is selected by the evaluation function Eval. Candidate sets from different inputs differ from each other at least along the dimensions that faithfulness constraints are sensitive to (for example, correspondence relations). Thus, a constraint  $\mathbb{C}$ 's assessment of any given output candidate *cand* may depend on *cand*'s relation to its input  $i$ , especially so if  $\mathbb{C}$  is a faithfulness constraint.

---

<sup>47</sup> By "logical necessity" in the passage quoted in the text, we might assume that Piggott means to convey some understanding of the consequences of actual PTC: since general  $\gg$  special deactivates special, there can be no evidence for special's existence, and since not all constraints are present in every language in Piggott's P&P framework, lack of evidence for the existence of a constraint constitutes evidence for that constraint's nonexistence. However, as discussed in §5.2 below (see especially footnote 51), PTC does *not* in fact guarantee that special is deactivated *tout court* under the general  $\gg$  special ranking.

Now consider the role of the constraint hierarchy. Whether  $C$ 's assessment of *cand* matters in practical terms — specifically, whether  $C$  crucially distinguishes *cand* from other output candidates *cand'* — depends on whether the members of *cand'* are all viable given the constraints that dominate  $C$ : if they are already ruled out by the portion of the constraint hierarchy  $CH$  ranked higher than  $C$ , then whatever  $C$  has to say about *cand* as opposed to *cand'* is rendered irrelevant, given the strictness of strict domination.

We can now appreciate Prince & Smolensky's definition of constraint activity, given in (50) below, accompanied by Prince & Smolensky's own prose clarification. Note that "separates the candidates" means "distinguishes some candidates from others" and that "admitted" means "not ruled out by".

(50) Constraint activity (Prince & Smolensky 1993/2004, §5.3)

- a. Definition: Let  $C$  be a constraint in a constraint hierarchy  $CH$  and let  $i$  be an input.  $C$  is *active on  $i$  in  $CH$*  if  $C$  separates the candidates in  $Gen(i)$  which are admitted by the portion of  $CH$  which dominates  $C$ .
- b. Prose clarification: "In other words, the portion of  $CH$  which dominates  $C$  filters the set of candidate parses of  $i$  to some degree, and then  $C$  filters it further. When  $C$  is not active for an input  $i$  in  $CH$ , the result of parsing  $i$  is not at all affected by the presence of  $C$  in the hierarchy."

With this background out of the way, we can better follow Prince & Smolensky's discussion leading up to the statement of PTC.

One consequence of the definition of harmonic ordering [= how constraints interact to select optimal forms—EB] is that there are conditions under which the presence of a more general constraint in a superordinate position in a hierarchy will eliminate all opportunities for a more specialized constraint in a subordinate position to have any effects in the grammar.

Note the ranking presupposed here: "a general constraint in a superordinate position" and "a more specialized constraint in a subordinate position". PTC does not dictate either order of the constraints, only that one order of the constraints (general above special, or  $G \gg S$ ) will, under certain conditions, lead to deactivation of  $S$ . These conditions are clarified in this "oversimplified first cut at the true result" of the theorem:

[I]f one constraint is more general than another in the sense that the set of inputs to which one constraint applies nonvacuously includes the other's nonvacuous input set, and if the two constraints *conflict* on inputs to which the more specific applies nonvacuously, then the more specific constraint must dominate the more general one in order for its effects to be visible in the grammar.

As clarified further below, a constraint  $C$  applies nonvacuously to an input  $i$  if  $C$  distinguishes among the members of  $i$ 's candidate output set, some members being better performers on  $C$  than others. If candidate sets are very inclusive, as much work in OT has suggested, then it is hard to conceive of any sensible constraints that *wouldn't* distinguish among the members of  $i$ 's entire candidate set. This issue, however, only impacts the broadest possible situation for which PTC is relevant: the deactivation of  $S$  due to its rank below  $G$  *in entire grammars*. In the context of 'elsewhere' effects, the precise conditions of interest apply at a level of analysis that corresponds to what we might intuitively call *process interaction* — overlap between

otherwise independently-motivated input-output disparities that we might analyze as separate rules in SPE. At this level of analysis, what matters is whether  $\mathbb{S}$  and  $\mathbb{G}$  conflict with each other over candidates that have not been rejected by the portion of the constraint hierarchy  $\mathbb{CH}$  that dominates  $\mathbb{S}$  and  $\mathbb{G}$ ; that is, situations in which  $\mathbb{S}$  and  $\mathbb{G}$  are at least potentially active in the sense of (50).

When  $\mathbb{CH}$  is the entire constraint hierarchy of a grammar, it is normally the case that [...] the constraint set is sufficient to winnow the candidate set down to a single output. [W]e will need to consider, more generally, initial portions of the constraint hierarchy of a grammar, *i.e.*, all the constraints from the highest-ranked down to some constraint which may not be the lowest-ranked. In these cases,  $\mathbb{CH}$  will often consist of just a few constraints, insufficient to winnow the candidate set down to a single parse; in that case,  $\mathbb{CH}$  will accept an entire set of parses, all equally harmonic, and all more harmonic than the competitors filtered out by  $\mathbb{CH}$ .

This leads to Prince & Smolensky's statement of PTC in (51) below, including definitions of two terms that are crucial to understanding the statement of the theorem.

(51) Pāṇini's Theorem on Constraint-ranking (PTC)

- a. *Nonvacuous application.* A constraint  $\mathbb{C}$  applies nonvacuously to an input  $i$  if  $\mathbb{C}$  separates  $\text{Gen}(i)$ , the set of candidate outputs of  $i$  admitted by UG.
- b. *Pāṇinian constraint relation.* Let  $\mathbb{S}$  and  $\mathbb{G}$  be two constraints.  $\mathbb{S}$  stands to  $\mathbb{G}$  as special to general in a Pāṇinian relation if, for any input  $i$  to which  $\mathbb{S}$  applies nonvacuously, any candidate output of  $i$  which satisfies  $\mathbb{S}$  fails  $\mathbb{G}$ .
- c. *Pāṇini's Theorem.* Let  $\mathbb{S}$  and  $\mathbb{G}$  stand as specific to general in a Pāṇinian relation. Suppose these constraints are part of a constraint hierarchy  $\mathbb{CH}$ , and that  $\mathbb{G}$  is active in  $\mathbb{CH}$  on some input  $i$ . Then, if  $\mathbb{G} \gg \mathbb{S}$ ,  $\mathbb{S}$  is not active on  $i$ .

### 5.1. Nonvacuous application and relativized candidate sets

The candidate set for every input is usually taken to be very inclusive in OT, with most if not all apparent limitations on the set being defined by the constraint ranking; this is the *freedom of analysis* property assumed (even if only implicitly) in most OT work (see McCarthy 2002:8-10). The definition of nonvacuous constraint application in (51)a may seem a little curious, then: depending on the exact extent of inclusiveness, virtually every constraint technically applies nonvacuously to every input.<sup>48</sup> Like constraint activity, nonvacuous application should be understood relative to a constraint hierarchy: we are concerned with the relevant residue of  $\text{Gen}(i)$  that is left once a particular irrelevant subset of  $\text{Gen}(i)$  is independently removed from consideration.

A schematic example will help to clarify. Suppose that the universal constraint set *Con* contains all and only the markedness constraints ONSET and NOCODA and

<sup>48</sup> This somewhat subtle point may have escaped Prince & Smolensky's notice because they do not specifically discuss inclusiveness, which was formally introduced by McCarthy & Prince (1993).

There are of course alternative architectures for OT (see McCarthy 2002:§3.3.3). Stratal architectures and less inclusive candidate set generator models (Wilson 2004, McCarthy, to appear), for example, both make it more likely that sensible constraints might not separate a given candidate set.

the faithfulness constraint DEP (the latter understood to be equivalent to Prince & Smolensky's FILL). Imagine now four inputs, /CV/, /CVC/, /V/, and /VC/, and suppose that the candidate sets for each input inclusively contains all possible syllabifications of the input, including all possible insertions of segments. If insertion is in principle unbounded, then of course these candidate sets are unbounded in size; however, there is a definable point at which insertion does not 'help' by offsetting a violation of a markedness constraint.

This point differs for each of the four inputs. For the input /CV/ it is immediate: insertion of any segment will provoke a violation of DEP with no improvement on either of the markedness constraints ONSET and NOCODA. (Indeed, most imaginable insertions will provoke further violation of one or both of the markedness constraints.) For the inputs /CVC/ and /V/ it is further along: insertion of a  $\mathbb{V}$  after the second C of /CVC/ can offset a violation of NOCODA, and insertion of a  $\mathbb{C}$  before the input V of /V/ can offset a violation of ONSET. For /VC/ it is further along still: a violation of NOCODA can be offset in the same way as for /CVC/ and a violation of ONSET can be offset in the same way as for /V/ — and, of course, a candidate with both insertions can simultaneously offset violations of both markedness constraints. (See Prince & Smolensky 1993/2004:§6.2.3.)

Candidates that go beyond this point in each case — those that exceed the number of 'helpful' insertions — are *harmonically bound*, which means that they cannot be optimal under any ranking of the constraints.<sup>49</sup> Thus, the pool of 'potential optima' for each input differs: for /CV/ it is just { [.CV.] }, while for /CVC/ it is { [.CVC.], [.CV.CV.] } and for /VC/ it is { [.VC.], [.V.CV.], [.CVC.], [.CV.CV.] }.<sup>50</sup> If we limit the definition of nonvacuous application to these candidate sets of possible optima, then the following table indicates how each constraint applies to each input (✓ = applies nonvacuously).

(52) Nonvacuous application relative to sets of potential optima

<i>Candidate sets of potential optima</i>	ONSET	NOCODA	DEP
a. /CV/ → { [.CV.] }			
b. /CVC/ → { [.CVC.], [.CV.CV.] }		✓	✓
c. /V/ → { [.V.], [.CV.] }	✓		✓
d. /VC/ → { [.VC.], [.V.CV.], [.CVC.], [.CV.CV.] }	✓	✓	✓

Because the candidate set of potential optima in (52)a consists of a single output, none of the constraints can separate the set and thus none of them applies nonvacuously in this case. DEP applies nonvacuously to the remaining, nonsingleton candidate sets of potential optima in (52)b-d, because DEP distinguishes candidates with insertion from those without (and candidates with  $n$  insertions from candidates with  $n+1$  insertions). ONSET applies nonvacuously to the candidate sets of potential optima in (52)c,d but not (52)b, since the former include onsetless syllables while the latter does not. Finally, NOCODA applies nonvacuously to the candidate sets of po-

<sup>49</sup> See Samek-Lodovici & Prince (1999) for extensive discussion of harmonic bounding.

<sup>50</sup> Note that I have also excluded harmonically bounded syllabifications (such as /CV/ → [.C.V.]) from consideration; these of course suffer the same fate as candidates with unhelpful insertion.

tential optima in (52)b,d but not (52)c, since the former include codaful syllables while the latter does not.

The point of this exercise is that the definition of nonvacuous application can be relativized to some well-defined subset of the entire candidate set  $\text{Gen}(i)$  for any input  $i$ . Here we have considered the well-defined subset of potential optima, but there are other subsets that can be useful in the determination of nonvacuous application and thereby in the identification of constraints standing in a Pāṇinian constraint relation.

## 5.2. Pāṇinian constraint relation and constraint deactivation

The central element of PTC is the definition of a Pāṇinian constraint relation in (51)b. Again, an example will clarify. To begin, take our example from §5.1 above, removing ONSET from *Con* and adding FINAL-C (McCarthy 1993). Also remove the inputs /V/ and /VC/ and add the inputs /CVCV/, /CVCVC/, /CVCCV/, and /CVCCVC/. The candidate sets of potential optima for all the inputs are as shown in (53). NOCODA and DEP now apply nonvacuously to all sets of potential optima, as does FINAL-C.

(53) Nonvacuous application relative to sets of potential optima

<i>Candidate sets of potential optima</i>	NOCODA	DEP	FINAL-C
a. /CV/ → {[.CV.], [.CVC.]}	✓	✓	✓
b. /VC/ → {[.CVC.], [.CV.CV.]}	✓	✓	✓
c. /CVCV/ → {[.CV.CV.], [.CV.CVC.]}	✓	✓	✓
d. /CVCVC/ → {[.CV.CVC.], [.CV.CV.CV.]}	✓	✓	✓
e. /CVCCV/ → {[.CVC.CV.], [.CVC.CVC.], [.CV.CV.CVC.]}	✓	✓	✓
f. /CVCCVC/ → {[.CVC.CVC.], [.CVC.CV.CV.], [.CV.CV.CVC.], [.CV.CV.CV.CV.]}	✓	✓	✓

According to the definition in (51)b, a constraint  $\mathbb{S}$  stands to another constraint  $\mathbb{G}$  as special to general in a Pāṇinian relation if, for any input  $i$  to which  $\mathbb{S}$  applies nonvacuously, any candidate output of  $i$  which satisfies  $\mathbb{S}$  fails (= violates)  $\mathbb{G}$ . Armed with this definition, we can see which pairs of constraints in (53), if any, stand as special to general in a Pāṇinian relation. Since a Pāṇinian relation must hold for *any input* to which  $\mathbb{S}$  applies nonvacuously, it suffices to find just one input for which this does not hold to establish that two constraints do not stand as special to general in a Pāṇinian relation.

Consider first NOCODA and its relation to DEP. In the case of /CV/, NOCODA is only satisfied by the output [.CV.], which also satisfies DEP. Therefore, no matter which is special and which is general, NOCODA and DEP do not stand in a Pāṇinian relation. Now consider FINAL-C and its relation to DEP. In the case of /CVCVC/, FINAL-C is only satisfied by the output [.CV.CVC.], which also satisfies DEP. Therefore, no matter which is special and which is general, FINAL-C and DEP also do not stand in a Pāṇinian relation.

Now consider FINAL-C and NOCODA. In all cases, FINAL-C is only satisfied by consonant-final candidates, all of which violate NOCODA. Therefore, FINAL-C and NOCODA stand as special to general (respectively) in a Pāṇinian relation. Thus, according to (51)c, PTC says that if NOCODA is active on some input  $i$  and NOCODA  $\gg$



FINAL-C, then FINAL-C is not active on *i*.<sup>51</sup> This corresponds with our intuitions about these constraints: if codas are not allowed, then word-final consonants are *a fortiori* not allowed.

Note that not all vowel-final candidates, which violate FINAL-C, satisfy NOCODA; for example, /CVCCV/ → [.CVC.CV.] violates both constraints. Therefore, NOCODA and FINAL-C do not stand as special to general (respectively) in a Pāṇinian relation. There is a reason why this reversal is worth considering. Imagine two constraints that are exact opposites of each other: NOCODA and YESCODA, or perhaps more plausibly FINAL-C and FINAL-V. Neither constraint in each pair is more general than the other, but they technically stand in a Pāṇinian relation nonetheless: for any input to which either constraint applies nonvacuously, any candidate output satisfying that constraint necessarily violates the other. According to (51)c, then, PTC says that whichever of the two constraints in each pair is lower-ranked, if the higher-ranked constraint is active on *i* then the lower-ranked one is not active on *i*. Pāṇinian relations technically need not be one-way.<sup>52</sup>

### 5.3. PTC and elsewhere conflict

Now consider the relationship between the more specific markedness constraint \$ and the more general, conflicting markedness constraint G in the OT analyses of the two types of complementary distribution in §3. As it turns out, \$ and G in these cases do not technically meet the conditions of the definition of a Pāṇinian constraint relation, and G can thus not be guaranteed to deactivate \$ — at least, not on its own — when  $G \gg \$$ . The problem is that there can be candidates in a candidate set that satisfy both \$ and G; however, this problem can be put to rest once the relevant subset of the candidate set is appropriately defined (as clarified in §5.1 above, and as I noted in the discussion in §3).

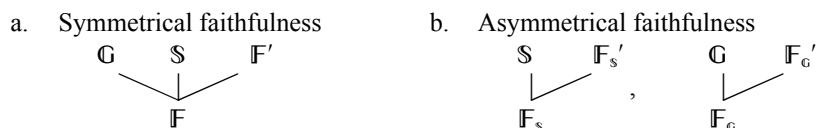
Take for example \*NV<sub>o</sub> (as \$) and \*V<sub>N</sub> (as G). Given an input to which \*NV<sub>o</sub> applies nonvacuously, such as /NV<sub>o</sub>/, there are candidate outputs, such as [OV<sub>o</sub>], which satisfy both constraints. As discussed in §3.1, these candidates violate some OTHER constraints, and it was assumed that those OTHER constraints are highly-enough ranked to give \*NV<sub>o</sub> and \*V<sub>N</sub> a chance to duke it out over the relevant candidates, [NV<sub>N</sub>] and [NV<sub>o</sub>]. (The same applies, of course, to \*[-high]<sub>s</sub> / (̀ Ci)V as \$ and \*V<sub>L</sub> / (̀ σ) as G; as noted in footnote 25, candidates with different foot-parsings that thereby satisfy both constraints must be ruled out by some appropriately highly-ranked constraints.)

<sup>51</sup> The proviso “if NOCODA is active on some input *i*” is meant to set aside situations in which NOCODA is itself deactivated by some higher-ranked constraint. Suppose, for example, that NOCODA is (only) dominated by some constraint X penalizing epenthesis at the right edges of, say, verbs. In nonverbs, X is irrelevant and NOCODA  $\gg$  FINAL-C guarantees the optimality of vowel-final words (= deactivation of FINAL-C). In verbs, on the other hand, X deactivates NOCODA and any decision between vowel-final and consonant-final words (accomplished via, say, vowel deletion) is free to be made by FINAL-C. (This is a case of the general *Emergence of the Unmarked* schema (McCarthy & Prince 1994).)

<sup>52</sup> This is an example of OT’s *total deactivation property* (Prince 1997b, 1999, 2000): given two constraints with the same violation classes — i.e., constraints that separate the candidate set into the same relational subsets, whether opposite or identical — only the higher-ranked of the two may be active.

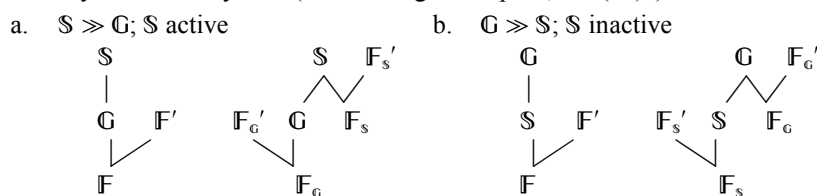
What is crucial is the assumption that the candidate set is appropriately narrowed down to the relevant subset of cases, those in which  $\$$  and  $\mathbb{G}$  compete head-to-head.<sup>53</sup> In order to properly establish this elsewhere conflict,  $\$$  and  $\mathbb{G}$  must be resolvable only by incompatible changes: nasalization and oralization, lengthening and shortening, insertion and deletion, and so on. If these incompatible changes are governed by a single, symmetrical faithfulness constraint  $\mathbb{F}$  (in the way that e.g. IDENT(nas) governs nasalization and oralization and that IDENT(long) governs lengthening and shortening), then  $\mathbb{F}$  must be ranked below both  $\$$  and  $\mathbb{G}$  and also below all other faithfulness constraints  $\mathbb{F}'$  governing any other changes that might also satisfy  $\$$  or  $\mathbb{G}$ .<sup>54</sup> If the incompatible changes are governed by separate, asymmetrical faithfulness constraints  $\mathbb{F}_s$  and  $\mathbb{F}_g$  (in the way that e.g. DEP governs insertion and MAX governs deletion), then  $\mathbb{F}_s$  must be ranked below  $\$$  and all other faithfulness constraints  $\mathbb{F}_s'$  governing any other changes that might also satisfy  $\$$ , and  $\mathbb{F}_g$  must be ranked below  $\mathbb{G}$  and all other faithfulness constraints  $\mathbb{F}_g'$  governing any other changes that might also satisfy  $\mathbb{G}$ . These ranking schemas are given in (54).

(54) Ranking schemas ensuring elsewhere conflict of  $\$$  and  $\mathbb{G}$



Under either of these ranking schemas, and restricting attention to the candidate subsets left behind by  $\mathbb{F}'$  (in the case of (54)a) or by  $\mathbb{F}_s'$  and  $\mathbb{F}_g'$  (in the case of (54)b), the markedness constraints  $\$$  and  $\mathbb{G}$  stand as special to general in a Pāṇinian constraint relation. According to PTC, then, if  $\$ \gg \mathbb{G}$  in a ranking consistent with one of the two schemas above, as shown in (55)a below, then  $\$$  may be active; if  $\mathbb{G} \gg \$$  in a ranking consistent with one of the two schemas above, as shown in (55)b below, then  $\$$  is inactive.

(55) Activity vs. inactivity of  $\$$  (relative to given input  $i$ ; see (51)c)



<sup>53</sup> I thank Alan Prince for helpful discussion of these ideas. I have incorporated most if not all of his suggestions, but I may not have done so in a way that coincides with his take on the topic.

<sup>54</sup> I make an important simplification here, common to most any discussion of ranking schemas of this type: that there are no other conflicting markedness constraints of interest. Retreat from this idealization in this case requires taking at least the following two species of markedness constraint into account:

- (i) A markedness constraint  $\mathbb{M}_1$  violated by the change governed by  $\mathbb{F}$ .
- (ii) A markedness constraint  $\mathbb{M}_2$  violated by one of the changes governed by  $\mathbb{F}'$ .

$\mathbb{M}_1$  must be ranked at least as low as its counterpart  $\mathbb{F}$ . On the other hand, only one or the other of  $\mathbb{M}_2$  and its counterpart in  $\mathbb{F}'$  needs to be ranked as highly as  $\mathbb{F}'$  is assumed to be in the discussion in the text.

Note that this is the same point made in the contemplation of the counterfactual ranking  $G \gg S$  at the end of each of the OT analyses in §3.1 and §3.2.

The proof of PTC, from the Appendix of Prince & Smolensky (1993/2004), is reproduced in the box below. The reader is invited to check the claims made in this section against this proof to verify that, indeed, PTC applies to two constraints  $S$  and  $G$  standing as special to general in a Pāṇinian constraint relation of the ‘elsewhere’ type.

**A.3 Pāṇini's Theorem on Constraint-ranking**

**Lemma.** Suppose a form  $f$  is accepted by a constraint hierarchy  $CH$  which includes  $C$ . Then either:  $f$  satisfies  $C$ , or  $C$  is not active on the input  $i$

**Proof.** Consider the filtering of the candidate set at  $C$ ;  $f$  must pass this filter since it passes all of  $CH$ . Suppose  $f$  violates  $C$ . Then  $f$  will be filtered out by  $C$  unless  $C$  is violated by all the candidates which pass the filtering in  $CH$  prior to  $C$ , in which case  $C$  is not active on  $i$ . So if  $f$  violates  $C$  then  $C$  is not active on  $i$ . This is equivalent to the statement in the Lemma.

**(112) Pāṇini's Theorem on Constraint-ranking (PTC).** Let  $S$  and  $G$  stand as specific to general in a Pāṇinian constraint relation. Suppose these constraints are part of a constraint hierarchy  $CH$ , and that  $G$  is active in  $CH$  on some input  $i$ . Then if  $G \gg S$ ,  $S$  is not active on  $i$ .

**Proof.** Assume  $S$  so placed in the hierarchy; then any form which is subject to evaluation by  $S$  must have survived the filtering of the top portion of the hierarchy down to and including  $G$ ; call this subhierarchy  $T$ . Let  $i$  be an input for which  $G$  is active, which exists by hypothesis. Consider the candidate set  $Gen(i)$ . By the lemma, any parse  $f$  in  $Gen(i)$  which survives  $T$  must satisfy  $G$ . So every parse  $f$  which survives to be evaluated by  $S$  satisfies  $G$ . But by the contrapositive of the Pāṇinian relationship, satisfying  $G$  entails failing  $S$ . Since all the candidates which  $S$  gets to evaluate fail  $S$ ,  $S$  cannot filter any of them out, and  $S$  is therefore not active on the input  $i$ .

#### 5.4. Final note: a common misinterpretation of the PTC

PTC defines conditions under which the more specific constraint of two conflicting constraints is deactivated, nothing more. A common misconception is that PTC somehow applies to constraints that stand in an inclusion relation but that do not conflict with each other: constraints in a *stringency relation* (Prince 1997bc, 1999, 2000, 2001). Two prominent examples of constraint types that are related in this way are (i) positional faithfulness constraints and their general faithfulness counterparts (Beckman 1995, 1997, 1998) and (ii) local conjunctions and their conjuncts (Smolensky 1993, 1995, 1997).<sup>55</sup> Beckman proposes that positional faithfulness constraints universally dominate their general faithfulness counterparts and Smolensky proposes that local conjunctions universally dominate their conjuncts, but neither of these ranking metaconditions has been demonstrated to be empirically necessary.<sup>56</sup> More relevant to present concerns, the more specific of two constraints in a stringency re-

<sup>55</sup> The markedness scale-sensitive constraints defended by de Lacy (2002, 2006) are also in a stringency relation; see de Lacy and also Prince (1997c, 1999, 2001) for discussion.

<sup>56</sup> In two separate instances (Baković 2000a:219ff, 2000b:9ff), I make empirically crucial use of the ranking metacondition proposed by Smolensky for local conjunctions and their conjuncts, but I'm the first to say that these don't count as 'demonstrations' in the strict sense intended here.

lation is not necessarily deactivated when ranked below the more general; see Prince (1997bc, 1999, 2000, 2001) for extensive discussion.

## 6. Concluding remarks

In §2 of this paper I showed that ‘elsewhere’ effects are instances of two basic types of complementary distribution, bounded and unbounded. In §3, I explained how an OT account of ‘elsewhere’ effects works, and demonstrated how that analysis unifies the two types of complementary distribution. There are two ways to achieve the same unification in SPE: either with a conjunctive, general-before-special order of two (feature-changing) rules — the prior rule specification analysis in (4) — or with a disjunctive, special-blocks-general order, preferably to be predicted by the EC (§4).

In §2.2.2 and again in §4.3, I discussed the problems with the conjunctive analysis with reference to the literature on the Duke of York gambit, and in §4.2 I noted six separate stipulations that constitute much of the EC. These stipulations are problematic precisely insofar as they must be stipulated; perfectly logical alternatives to each stipulation have been discussed, proposed, and argued about in the literature on the EC and on disjunctive ordering more generally, with the only deciding factors being empirical. It is thus significant that the empirically correct form of disjunctive application falls out from OT’s most basic assumptions, as originally noted by Prince & Smolensky (1993/2004).

One must conclude, as I have, that all the commotion surrounding the EC in the literature is due to the intuitive correctness of its premise and the logical consequence of that premise. The premise is that elsewhere effects are not about derivational rule interactions but rather about output comparisons; its consequence is that elsewhere effects are not about the undoing of earlier-ordered general rules but rather about the violation of lower-ranked general constraints. As Prince (1997a:2) notes in response to Halle (1995):

Observe that the Elsewhere Condition determines which one of two or more competing processes will apply in a given environment, blocking the expected derivational relationship between them. [Halle’s] aim, then, is to argue from the *failure* of serial derivation to the conclusion that serial derivation is necessary: the logic is not going to be straightforward. By contrast, Optimality Theory is thoroughly in the business of selecting which (of many) competing input-output maps will prevail and needs no Elsewhere Condition to guide it. As Prince & Smolensky observe, Elsewhere Condition disjunctivity is a sub-case of the general mode of constraint interaction in OT.

### Encomium

I first met Alan soon after his 45<sup>th</sup> birthday, at the 1991 LSA Linguistic Institute at UC Santa Cruz, one of the first public events where Alan and Paul Smolensky introduced the fruits of their collaboration. I was just a wide-eyed undergraduate at the time and hadn't the slightest clue what was going on. A year and a half later, while visiting Rutgers (and after I'd learned a little about OT), I found Alan in the copy-room making and arranging stacks upon stacks of copies of Prince & Smolensky (1993) to send to linguists around the world. I was blown away by Alan's foresight (quite radical at the time): he quickly built up and maintained support for the theory by actually *sharing the ideas as immediately and as widely as possible*. I'm honored to have been a part of what Alan and Paul started, and I'm especially lucky to have had Alan as a teacher, advisor, and friend.

### References

- ROA = Rutgers Optimality Archive, <http://roa.rutgers.edu/>.
- Alderete, J. 1999. *Morphologically Governed Accent in Optimality Theory*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA-309.]
- Anderson, S. R. 1969. *West Scandinavian Vowel Systems and the Ordering of Phonological Rules*. Doctoral dissertation, MIT.
- Anderson, S. R. 1974. *The Organization of Phonology*. New York: Academic Press.
- Anderson, S. R. and W. Browne. 1973. On keeping exchange rules in Czech. *Papers in Linguistics* 6, 445-482.
- Bach, E. 1968. Two proposals concerning the simplicity metric in phonology. *Glossa* 2, 128-149.
- Baković, E. 1996. Foot harmony and quantitative adjustments. ROA-168.
- Baković, E. 1999. Deletion, insertion, and symmetrical identity. In B. Vaux and S. Kuno (eds.), *Harvard Working Papers in Linguistics* 7, 27-28. [ROA-300.]
- Baković, E. 2000a. *Harmony, Dominance and Control*. Doctoral dissertation, Rutgers University. [ROA-360.]
- Baković, E. 2000b. Nasal place neutralization in Spanish. *Penn Working Papers in Linguistics* 7.1, 1-12. [ROA-386.]
- Beckman, J. 1995. Shona height harmony: markedness and positional identity. In J. Beckman, L. Walsh Dickey, and S. Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18, 53-75. Amherst, MA: GLSA Publications.
- Beckman, J. 1997. Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14, 1-46.
- Beckman, J. 1998. *Positional Faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA-234.]
- Bever, T. 1967. *Leonard Bloomfield and the Phonology of the Menomini Language*. Doctoral dissertation, MIT.
- Burzio, L. 1994. *Principles of English Stress*. Cambridge: Cambridge University Press.
- Burzio, L. 1995. The Rise of Optimality Theory. *Glott International* 1.6.
- Burzio, L. 1996. A Reply to Prof. Halle. *Glott International* 2.4.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1968. *Language and Mind*. New York: Harcourt Brace Jovanovich.

- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English*. Cambridge, MA: MIT Press.
- Cohn, A. and J. McCarthy. 1994. Alignment and parallelism in Indonesian phonology. ROA-25. [Published 1998 in *Working Papers of the Cornell Phonetics Laboratory* 12, 53-137.]
- de Lacy, P. 2002. *The Formal Expression of Markedness*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA-542.]
- de Lacy, P. 2006. *Markedness: Reduction and Preservation*. Cambridge: Cambridge University Press.
- Goldsmith, J. 1981. Towards an autosegmental theory of accent: the case of Tonga. Distributed by the Indiana University Linguistics Club.
- Goldsmith, J. 1984a. Tone and accent in Tonga. In G. N. Clements and J. Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, 19-51. Cinnaminson: Foris.
- Goldsmith, J. 1984b. Meeussen's rule. In M. Aronoff and R. Oehrle (eds.), *Language Sound Structure*, 245-259. Cambridge, MA: MIT Press.
- Goldsmith, J. and B. Laks. To appear. Generative phonology: its origins, its principles, and its successors. In L. Waugh and J. E. Joseph (eds.), *The Cambridge History of Linguistics*. Cambridge: Cambridge University Press. [<http://hum.uchicago.edu/~jagoldsm/Papers/GenerativePhonology.pdf>.]
- Gouskova, M. 2003. *Deriving Economy: Syncope in Optimality Theory*. Doctoral dissertation, University of Massachusetts, Amherst. [ROA-610.]
- Green, A. 1996. Some effects of the Weight-to-Stress Principle and Grouping Harmony in the Goidelic languages. *Working Papers of the Phonetics Laboratory* 11, 117-155. [ROA-153.]
- Green, A. 2001. The tense-lax distinction in English vowels and the role of parochial and analogical constraints. *Linguistics in Potsdam* 16, 32-57. [ROA-464.]
- Halle, M. 1971. Remarks on Slavic accentology. *LI* 2, 1-20.
- Halle, M. 1977. Tenseness, vowel shift, and the phonology of back vowels in Modern English. *LI* 8, 611-625.
- Halle, M. 1978. Formal vs. functional considerations in phonology. *Studies in the Linguistic Sciences* 8.2, 123-134. [Also in B. Brogyanyi (ed.), *Studies in Diachronic, Synchronic, and Typological Linguistics: Festschrift for Oswald Szeмерényi on the Occasion of his 65th Birthday*, 325-341. Amsterdam: John Benjamins, 1979.]
- Halle, M. 1995. Comments on Burzio (1995). *Glott International* 1.9/10, 27-28.
- Halle, M. 2002. *From Memory to Speech and Back: Essays 1952-2002*. Berlin: Mouton de Gruyter.
- Halle, M. and W. Idsardi. 1997. *r*, Hypercorrection, and the Elsewhere Condition. In I. Roca (ed.), *Derivations and Constraints in Phonology*, 331-348. Oxford: Clarendon Press.
- Halle, M. and W. Idsardi. 1998. A response to Alan Prince's letter. *Glott International* 3.1, 1 & 22.
- Halle, M. and J.-R. Vergnaud. 1982. On the framework of Autosegmental Phonology. In H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representations, Pt. 1*, 65-82.

- Halle, M. and J.-R. Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.
- Harris, J. 1994. *English Sound Structure*. Oxford: Blackwell.
- Harris, J. W. 1969. *Spanish Phonology*. Cambridge, MA: MIT Press.
- Harris, J. W. 1974. Evidence from Portuguese for the ‘Elsewhere Condition’ in phonology. *LI* 5, 61-80.
- Hill, J. 1970. A peeking rule in Cupeño. *LI* 1, 534-539.
- Howard, I. 1972. *A Directional Theory of Rule Application in Phonology*. Doctoral dissertation, MIT.
- Howard, I. 1975. Can the ‘Elsewhere Condition’ get anywhere? *Language* 51, 109-127.
- Johnson, C. D. 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.
- Kahn, D. 1976. *Syllable-based generalizations in English phonology*. Doctoral dissertation, MIT.
- Kenstowicz, M. 1994a. *Phonology in Generative Grammar*. Cambridge: Blackwell.
- Kenstowicz, M. 1994b. On Metrical Constituents: Unbalanced Trochees and Degenerate Feet. In J. Cole and C. Kisseberth (eds.), *Perspectives in Phonology*, 113-131. Stanford, CA: CSLI Publications.
- Kiparsky, P. 1973. “Elsewhere” in phonology. In *A Festschrift for Morris Halle*, S. R. Anderson and P. Kiparsky (eds.), 93-106. New York: Holt, Reinhart, & Winston.
- Kiparsky, P. 1982. Lexical phonology and morphology. In *The Linguistics Society of Korea (ed.), Linguistics in the Morning Calm*, 3-91. Seoul: Hanshin Publishing Co.
- Kisseberth, C. 1970. On the functional unity of phonological rules. *LI* 1, 291-306.
- Koutsoudas, A., G. Sanders, and C. Noll. 1974. The application of phonological rules. *Language* 50, 1-28.
- Langacker, R. 1969. Mirror-image rules II: lexicon and phonology. *Language* 45, 844-862.
- McCarthy, J. 1991. Synchronic rule inversion. *BLS* 17, 192-207.
- McCarthy, J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 38, 169-195.
- McCarthy, J. 1999a. Sympathy and phonological opacity. *Phonology* 16, 331-399.
- McCarthy, J. 1999b. Appendix: a note on Boston *r* and the Elsewhere Condition. Available at <http://people.umass.edu/jjmccart/appendix.pdf>.
- McCarthy, J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, J. 2003. Sympathy, culminativity, and the Duke-of-York gambit. In C. Féry and R. van de Vijver (eds.), *The Syllable in Optimality Theory*, 23-76. Cambridge: Cambridge University Press.
- McCarthy, J. to appear. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox Publishers.
- McCarthy, J. and A. Prince. 1993. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Technical report RuCCS-TR-3, New Brunswick, NJ: Rutgers University Center for Cognitive Science. [ROA-482.]

- McCarthy, J. and A. Prince. 1994. The Emergence of the Unmarked: Optimality in prosodic morphology. *NELS* 24, 333-379. [ROA-13; also excerpted in J. McCarthy (ed.), *Optimality Theory in Phonology: A Reader*. Malden, MA: Blackwell, 2004.]
- McCarthy, J. and A. Prince. 1995. Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey, and S. Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18, 249-384. Amherst, MA: GLSA Publications. [ROA-103.]
- McCarthy, J. and A. Prince. 1999. Faithfulness and identity in Prosodic Morphology. In R. Kager, H. van der Hulst, and W. Zonneveld (eds.), *The Prosody-Morphology Interface* 218-309. Cambridge: Cambridge University Press. [ROA-216; also excerpted in J. McCarthy (ed.), *Optimality Theory in Phonology: A Reader*. Malden, MA: Blackwell, 2004.]
- Mester, R. A. 1994. The quantitative trochee in Latin. *NLLT* 12, 1-61.
- Moreton, E. 1996/1999. Non-computable functions in Optimality Theory. ROA-364. [Excerpted in J. McCarthy (ed.), *Optimality Theory in Phonology: A Reader*. Malden, MA: Blackwell, 2004.]
- Myers, S. 1987. Vowel shortening in English. *NLLT* 5, 485-518.
- Norton, R. J. 2003. *Derivational Phonology and Optimality Phonology: Formal Comparison and Synthesis*. Doctoral dissertation, University of Essex. [ROA-613.]
- Piggott, G. L. 2003. Theoretical implications of segment neutrality in nasal harmony. *Phonology* 20, 375-424.
- Prince, A. 1990. Quantitative consequences of rhythmic organization. *CLS* 26-II: *Papers from the Parasession on the Syllable in Phonetics and Phonology*, 355-398.
- Prince, A. 1993. Minimal violation. Handout of talk presented at the Rutgers Optimality Workshop (ROW-1), Rutgers University, New Brunswick.
- Prince, A. 1997a. Elsewhere and otherwise. ROA-217. [Shorter version appears in *Glott International* 2.1, 23-24.]
- Prince, A. 1997b. Stringency and Anti-Paninian Hierarchies. LSA Linguistic Institute class handout. [<http://ling.rutgers.edu/gamma/talks/insthdt2.pdf>.]
- Prince, A. 1997c. Paninian relations. Handout of talk presented at the University of Massachusetts, Amherst. [<http://ling.rutgers.edu/gamma/talks/umass1997.pdf>.]
- Prince, A. 1999. [Son of] Paninian relations. Handout of talk presented at the University of Marburg. [<http://ling.rutgers.edu/gamma/talks/panrel-mbg.pdf>.]
- Prince, A. 2000. The special and the general. Handout of invited talk presented at the 24<sup>th</sup> Penn Linguistics Colloquium. [<http://ling.rutgers.edu/gamma/talks/plc-hdt.pdf>.]
- Prince, A. 2001. Invariance under re-ranking. Handout of plenary talk presented at *WCCFL* 20. [<http://ling.rutgers.edu/gamma/talks/WCCFL2001.pdf>.]
- Prince, A. and P. Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Technical report RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science. [ROA-537. Published 2004, Malden, MA: Blackwell.]
- Pullum, G. 1976. The Duke of York gambit. *Journal of Linguistics* 12, 83-102.



- Samek-Lodovici, V. and A. Prince. 1999. Optima. Ms., University of London and Rutgers University, New Brunswick, NJ. [ROA-363.]
- Schane, S. 1984. Two English vowel movements: a particle analysis. In M. Aronoff and R. Oehrle (eds.), *Language Sound Structure*, 32-51. Cambridge, MA: MIT Press.
- Scobbie, J. 1991. *Attribute Value Phonology*. Doctoral dissertation, University of Edinburgh.
- Scobbie, J. 1993. Constraint violation and conflict from the perspective of Declarative Phonology. *Canadian Journal of Linguistics* 38, 154-168.
- Smolensky, P. 1993. Harmony, markedness, and phonological activity. Handout of talk presented at the Rutgers Optimality Workshop (ROW-1), Rutgers University, New Brunswick. [ROA-37.]
- Smolensky, P. 1995. On the internal structure of the constraint component Con of UG. Handout of talk presented at the University of California, Los Angeles. [ROA-86.]
- Smolensky, P. 1997. Constraint interaction in generative grammar II: local conjunction, or random rules in Universal Grammar. Handout of talk presented at the Hopkins Optimality Theory Workshop, Baltimore, MD.
- van de Vijver, R. 2003. The CiV-generalization in Dutch: what *petunia*, *mafia*, and *sovjet* tell us about Dutch syllable structure. In C. Féry and R. van de Vijver (eds.), *The Syllable in Optimality Theory*, 338-355. Cambridge: Cambridge University Press.
- Wilson, Colin. 2004. Analyzing unbounded spreading with constraints: marks, targets, and derivations. Ms., University of California, Los Angeles.

Linguistics Department  
9500 Gilman Dr. #0108  
University of California, San Diego  
La Jolla, CA 92093-0108  
[bakovic@ling.ucsd.edu](mailto:bakovic@ling.ucsd.edu)  
<http://ling.ucsd.edu/~bakovic/>