2 Opacity and Ordering

ERIC BAKOVIĆ

1 Introduction

Few notions in phonological theory have received as much attention in the literature as OPACITY. In the almost 40 years since Kiparsky (1971, 1976) offered the definition given in (1), the bulk of the attention paid to opacity has been relatively recent and has been fueled by the field's massive (but incomplete) shift from the rule-based serialism framework of *The Sound Pattern of English* (Chomsky and Halle 1968) to the constraint-based parallelism framework of Optimality Theory (Prince and Smolensky 1993).

(1) Opacity (Kiparsky 1976: 79)

A phonological rule $\mathbb{P}$ of the form $A \rightarrow B \ / \ C\_\_D$ is OPAQUE if there are surface structures with either of the following characteristics:

a. instances of A in the environment $C\_\_D$.
b. instances of B derived by $\mathbb{P}$ that occur in environments other than $C\_\_D$.

According to (1), the opacity of a (hypothesized) rule $\mathbb{P}$ can be formally diagnosed by comparing the set of (predicted) surface representations with the generalization expressed by $\mathbb{P}$: to say that $\mathbb{P}$ is opaque is to say that the applicability or application of $\mathbb{P}$ is (somehow) obscured on the surface. Kiparsky's substantive claim was that an opaque rule $\mathbb{P}$ is *difficult to learn*, either (1a) because there are surface counterexamples to $\mathbb{P}$'s applicability, or (1b) because there are surface contexts in which $\mathbb{P}$'s application is not motivated.

Kiparsky's support for this substantive learnability claim was a set of examples of language change in which previously opaque rules become transparent. More

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu

specifically, Kiparsky identified two pairwise orders between rules made possible by the rule-based serialism framework, COUNTERFEEDING and COUNTERBLEEDING, and argued that each order (i) results in a particular type of opacity ((1a) and (1b), respectively), and (ii) tends to change over time to the corresponding reverse, transparent order (FEEDING and BLEEDING, respectively). These orders (and Kiparsky's claims) are discussed more extensively in Section 2.

If there's only one thing that phonologists have learned from Kiparsky's work on the subject of opacity, it is to equate opacity of type (1a) with counterfeeding and opacity of type (1b) with counterbleeding. My aim here is to demonstrate that these equations are falsified in both directions: in Section 3 I show that not all cases of type (1a) opacity result from counterfeeding and that not all cases of counterfeeding result in opacity of either type, and in Section 4 I show that not all cases of type (1b) opacity result from counterbleeding and that not all cases of counterbleeding result in opacity of either type. This demonstration reveals a very different, more complex, and more complete picture of what opacity is than previously conceived. This is a significant result because opacity's original raison d'être is Kiparsky's claim that an opaque rule is difficult to learn. This claim is meaningful and testable only insofar as we have a clear understanding of what is and what is not an instance of an opaque rule, and what an account of such an instance, in turn, should look like.

2 Pairwise Rule Ordering

The central principle of rule-based serialism is rule ordering. Bromberger and Halle's (1989: 58–59) informal definition of rule ordering, given in (2), suffices for our purposes.

(2) Rule ordering (Bromberger and Halle 1989: 58–59)

Phonological rules are ordered with respect to one another. A phonological rule R does not apply necessarily to the underlying representation; rather, R applies to the derived representation that results from the application of each applicable rule preceding R in the order of the rules.

There are four recognized non-trivial pairwise ordered rule relations in rule-based serialism: feeding, bleeding, counterfeeding, and counterbleeding. These are defined informally in (3).[1]

(3) Pairwise ordered rule relations (adapted from McCarthy 2007b)

Given two rules $\mathbb{A}$, $\mathbb{B}$ such that $\mathbb{A}$ precedes $\mathbb{B}$,

a. $\mathbb{A}$ FEEDS $\mathbb{B}$ iff $\mathbb{A}$ creates additional inputs to $\mathbb{B}$.
b. $\mathbb{A}$ BLEEDS $\mathbb{B}$ iff $\mathbb{A}$ eliminates potential inputs to $\mathbb{B}$.
c. $\mathbb{B}$ COUNTERFEEDS $\mathbb{A}$ iff $\mathbb{B}$ creates additional inputs to $\mathbb{A}$.
d. $\mathbb{B}$ COUNTERBLEEDS $\mathbb{A}$ iff $\mathbb{B}$ eliminates potential inputs to $\mathbb{A}$.

Note that counterfeeding and counterbleeding are *counterfactual inverses* of feeding and bleeding, respectively, because counterfeeding *would be* feeding and counterbleeding *would be* bleeding if the two rules involved were ordered in the opposite way. The terminology, though notoriously difficult to learn, is thus not completely misleading.

Two rules may interact in different ways in different derivations. Consider (4), for example. In (4a), Deletion feeds Palatalization: deletion of the /u/ crucially places the preceding /t/ before a [–back] vowel. In (4b), on the other hand, Deletion bleeds Palatalization: the deleted /i/ is [–back] and thus would have induced palatalization of the preceding /t/ if it hadn't been deleted. In both (4c) and (4d), the two rules are mutually non-affecting: in (4c), neither vowel is [–back] and so the /t/ is never in a context to be palatalized; in (4d), both vowels are [–back] and so the /t/ is in a context to be palatalized either way.

(4) Feeding and bleeding in different derivations (hypothetical)

		a. /tue/	b. /tio/	c. /tou/	d. /tei/
Deletion:	V → ∅ / __ V	∅	∅	∅	∅
Palatalization:	t → ʧ / __ [–back]	ʧ			ʧ
		[ʧe]	[to]	[tu]	[ʧi]

Reversing the order of these two rules, as in (5), we get counterfeeding and counterbleeding in different derivations.[2] In (5a), Deletion counterfeeds Palatalization: deletion of the /u/ places the preceding /t/ before a [–back] vowel, but too late for Palatalization to do anything about it. In (5b), on the other hand, Deletion counterbleeds Palatalization: the deleted /i/ is [–back] and thus induces palatalization of the preceding /t/ before deleting. In both (5c) and (5d), the two rules are again mutually non-affecting, just as in (4) above.

(5) Counterfeeding and counterbleeding in different derivations (hypothetical)

		a. /tue/	b. /tio/	c. /tou/	d. /tei/
Palatalization:	t → ʧ / __ [–back]		ʧ		ʧ
Deletion:	V → ∅ / __ V	∅	∅	∅	∅
		[te]	[ʧo]	[tu]	[ʧi]

Although (3) constitutes a useful picture of the typology of possible ordered rule relations predicted by the central principle of rule-based serialism in (2), it is still defined (almost) exclusively in terms of interactions between just two ordered rules. I hardly hesitate to qualify this statement because most if not all definitions of pairwise ordered rule relations provided in textbooks and in the scholarly literature are insufficiently precise about situations involving more than two rules, which may counterintuitively fit or not fit a given definition.[3] But the fact remains that the bulk of the relevant literature focuses on pairwise interactions.

There have been two significant proposals for classifying the ordering relations in (3). The first was the relatively formal hypothesis that "rules tend to shift into the order which allows their fullest utilization in the grammar" (Kiparsky 1968c:

200). This privileges feeding and counterbleeding orders, grouping them together as "unmarked" because these are the orders in which both rules apply non-vacuously – that is, in which the two rules are both utilized, as can be appreciated from the feeding derivation in (4a) and the counterbleeding derivation in (5b) above. Conversely, bleeding and counterfeeding orders are "marked" because these are the orders in which one of the two rules fails to apply non-vacuously, as can be appreciated from the bleeding derivation in (4b) and the counterfeeding derivation in (5a).

There were several challenges to Kiparsky's "maximal utilization" hypothesis; see Ken stowicz and Kisseberth (1977: 159ff.) for an informative summary critique. Kiparsky's response was a relatively substantive second hypothesis, that "rules tend to be ordered so as to become maximally transparent" (Kiparsky 1971: 623). A transparent rule is one that does not meet either of the two conditions defined in (1) above, repeated in (6) below.

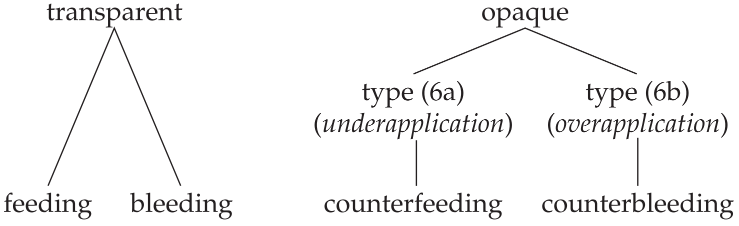
(6) Opacity, repeated from (1)

A phonological rule $\mathbb{P}$ of the form $A \rightarrow B \,/\, C\_\_D$ is opaque if there are surface structures with any of the following characteristics:
a. instances of A in the environment $C\_\_D$.
b. instances of B derived by $\mathbb{P}$ that occur in environments other than $C\_\_D$.

Kiparsky hypothesized that diachronic change proceeds from harder-to-learn opacity-promoting rule orders to easier-to-learn transparency-promoting ones, modulo potentially conflicting principles such as paradigm uniformity. Kaye (1974, 1975), Kisseberth (1976), and Kenstowicz and Kisseberth (1977: 170ff.) question the overall learnability claim by pointing out that phonological opacity often helps to maintain lexical contrasts (which one might think of as "semantic transparency"); see Łubowicz (2003a) for a recent rearticulation of this view.[4]

McCarthy (1999) adapts a couple of terms from work on reduplication by Wilbur (1973), underapplication and overapplication, to elucidate the two types of opacity in (6).[5] Type (6a) describes situations in which there are surface representations to which $\mathbb{P}$ could apply non-vacuously; $\mathbb{P}$ has thus *underapplied*. Type (6b) describes situations in which there are surface representations to which $\mathbb{P}$ has applied non-vacuously, but which do not otherwise meet $\mathbb{P}$'s structural description; $\mathbb{P}$ has thus *overapplied*. Kiparsky's explicit and subsequently generally accepted classification of the four pairwise rule interactions in (3) is shown in (7).

(7) Classification of pairwise ordered rule interactions (Kiparsky 1971, 1976)



In the next two sections I demonstrate that the classification of pairwise ordered rule interactions, in (7) is misleading at best. Counterfeeding is but one of several devices that can be and have been used to describe actual examples meeting the definition of underapplication in (6a), and counterfeeding does not always lead to underapplication (Section 3). Similarly, counterbleeding is not the only way to describe actual examples meeting the definition of overapplication in (6b), and counterbleeding does not always lead to overapplication (Section 4).

3 Underapplication and Counterfeeding

The definitions of underapplication opacity in (6a) and of the counterfeeding relation in (3c) are repeated (in suitably modified forms) in (8) and (9), respectively.

(8) A phonological rule $\mathbb{P}$ of the form $A \rightarrow B$ / $C\_\_D$ UNDERAPPLIES if there are surface structures with instances of A in the environment $C\_\_D$.

(9) $\mathbb{B}$ COUNTERFEEDS $\mathbb{A}$ iff $\mathbb{B}$ creates additional inputs to $\mathbb{A}$ and $\mathbb{A}$ precedes $\mathbb{B}$.

I begin in Section 3.1 by explaining how some examples of counterfeeding as defined in (9) result in underapplication as defined in (8). Then I demonstrate that counterfeeding is not the only source of underapplication. In Section 3.2 I discuss various types of BLOCKING, the most obvious type of underapplication that is not typically categorized as such in the literature, and in Section 3.3 I discuss a handful of other phenomena that also arguably contribute to underapplication opacity: the restriction of a rule to particular lexical classes or levels, rule exceptions, and rule optionality. Finally, I demonstrate in Section 3.4 that counterfeeding does not always lead to underapplication opacity, at least not as underapplication is defined in (8).

3.1 Counterfeeding

The counterfeeding relation in (9) describes situations where a later-ordered rule $\mathbb{B}$ creates representations to which an earlier-ordered rule $\mathbb{A}$ could have applied non-vacuously; modulo the action of other, even later rules (see Section 3.4), $\mathbb{A}$ underapplies in such situations. This was exemplified by the hypothetical derivation of /tue/ in (5a) above: Deletion creates an additional input to Palatalization, but because Palatalization precedes Deletion the result is a surface structure, [te], with a voiceless coronal stop before a front vowel – the structural description of Palatalization. Palatalization has thus underapplied in this derivation.

Following McCarthy (1999), I distinguish COUNTERFEEDING ON ENVIRONMENT from COUNTERFEEDING ON FOCUS interactions (see also Baković 2007: 221ff.). In a rule of the form $A \rightarrow B$ / $C\_\_D$, the *focus* is A, the element to be changed by the rule, and the *environment* is $C\_\_D$, the necessary context surrounding the focus. In counterfeeding on environment interactions the later-ordered rule $\mathbb{B}$ creates the environment of the earlier-ordered rule 𝔸, and in counterfeeding on focus interactions 𝔹 creates the focus of 𝔸. The main significance of this distinction is that cases of counterfeeding on focus have comparably successful accounts without ordering, as will be briefly noted in Section 3.1.2.

3.1.1 Counterfeeding on Environment

Consider as an example of counterfeeding on environment the following two rules of Lomongo.

(10) Counterfeeding in Lomongo (Hulstaert 1961; Kenstowicz and Kisseberth 1979)

	a. /o+bina/	b. /o+isa/	c. /ba+bina/
Gliding: [−low] → [−syll] / __V		w	
Deletion: $\begin{bmatrix}+\text{voi}\\ -\text{son}\end{bmatrix} \to \emptyset$ / V __	Ø		Ø
	[o+ina]	[w+isa]	[ba+ina]

Glosses: (10a) ‘you (sg.) dance’, (10b) ‘you (sg.) hide’, (10c) ‘they dance’

The derivations in (10b–c) illustrate the independent action of each of the rules: gliding applies alone in (10b) and Deletion applies alone in (10c), with no interaction in either case. In (10a), Deletion counterfeeds Gliding by creating the environment (a following vowel) that Gliding could have used to apply to the /o/. Gliding thus underapplies because there are surface representations with non-low prevocalic vowels that have not become glides.

There are also more complex interactions involving counterfeeding on environment, for instance where 𝔸 feeds 𝔹 but 𝔹 in turn counterfeeds 𝔸. I borrow from Kavitskaya and Staroverov (2010) the term “fed counterfeeding” to refer to this type of interaction. An example of FED COUNTERFEEDING ON ENVIRONMENT is found in Lardil, as shown in (11).

(11) Fed counterfeeding in Lardil (Hale 1973; Kavitskaya and Staroverov 2010)[6]

	a. /dibirdibi/	b. /yiliyili/	c. /wangalk/
Apocope: V → Ø / $\sigma\,\sigma$ __ #	Ø	Ø	
Deletion: [−apical] → Ø / __ #	Ø		Ø
	[dibirdi]	[yiliyil]	[wangal]

Glosses: (11a) ‘rock cod’, (11b) ‘oyster species’, (11c) ‘boomerang’

The derivations in (11b–c) again illustrate the independent action of each of the rules: in (11b), application of Apocope leaves a word-final apical consonant behind, which is not subject to Deletion; in (11c), there is no word-final vowel before or after application of Deletion. In (11a), Apocope feeds Deletion: removal of the word-final vowel places the preceding non-apical consonant in a position to be deleted. But Deletion also counterfeeds Apocope here: deletion of the non-apical consonant places the preceding vowel in a position to also be removed by Apocope, but Apocope does not apply to this vowel. Apocope thus underapplies because there are surface representations with word-final vowels.

3.1.2 Counterfeeding on Focus Now consider as an example of counterfeeding on focus the following rules of Western Basque.

(12) Counterfeeding in Western Basque (de Rijk 1970; Hualde 1991; Kawahara 2002)

		a. /alaba+a/	b. /seme+e/
Raising-to-High:	[−low] → [+high] / __V		i
Raising-to-Mid:	[+low] → [−low] / __V	e	
		[alabe+a]	[semi+e]

Glosses: (12a) ‘daughter’, (12b) ‘son’

The derivation in (12b) illustrates the independent action of Raising-to-High, which applies alone here to raise the prevocalic mid vowel. In (12a), Raising-to-Mid applies and counterfeeds Raising-to-High by changing the focus to a mid vowel that Raising-to-High could have applied to if it were later in the order. Raising-to-High thus underapplies because there are surface representations with mid prevocalic vowels that have not become high.

Examples of counterfeeding on focus like this one, particularly when the environments of the rules are the same, are referred to as CHAIN SHIFTS: underlying *A* becomes *B* and underlying *B* becomes *C*, but an *A* that becomes a *B* does not go on to become a *C*.

A comparably successful alternative to the ordering analysis of chain shifts recognizes the scalar nature of the dimensions along which chain shifts tend to occur (Kirchner 1996; Baković 1996; Gnanadesikan 1997; Kawahara 2002; Moreton and Smolensky 2002): movement toward the target end of the scale, even if it is not all the way, is better than no movement at all. In Western Basque, for example, the relevant scale is that of vowel height and the target end of the scale is a high vowel; both underlying mid and underlying low vowels aim in the right direction, though only mid vowels manage to hit the target.

Another comparably successful alternative capitalizes on the fact that chain shifts are *contrast-preserving* (Łubowicz 2003a, b): the fact that underlying *A* surfaces as *B* and underlying *B* surfaces as *C* means that the underlying contrast between *A* and *B* is manifested as a contrast, albeit a shifted one, on the surface. (See Łubowicz-Baković 2011, and references therein for more details on chain shifts and their analysis.)

There are also examples of FED COUNTERFEEDING ON FOCUS, for example in Nootka:

(13) Fed counterfeeding in Nootka (Sapir and Swadesh 1978; McCarthy 1999, 2003, 2007b)

	a. /muː q/	b. /ħaju+q i/	c. /ɬaː kʷ+ʃit͡ɬ/
Labialization:			
[+dors] → [+rnd] / [+rnd] __	qʷ	qʷ	
Delabialization:			
[+dors] → [−rnd] / __ $]_\sigma$	q		k
	[muː q]	[ħaju+qʷi]	[ɬaːk +ʃit͡ɬ]

Glosses: (13a) ‘throwing off sparks’, (13b) ‘ten on top’, (13c) ‘to take pity on’

The derivations in (13b–c) yet again illustrate the independent action of each of the rules. In (13b), Labialization applies to a dorsal that is preceded by a round vowel but is not syllable-final, and so Delabialization is inapplicable; in (13c), Delabialization applies to a syllable-final dorsal that is not preceded by a round vowel, and so Labialization is inapplicable. In (13a), Labialization both feeds and is counterfed by Delabialization: the dorsal is preceded by a round vowel and so it labializes, but this creates a syllable-final labialized dorsal that is subsequently delabialized – which puts the dorsal back in the position of being non-vacuously subject to Labialization. Thus, even though Labialization "applies" in the sense that it makes a non-vacuous change during the course of the derivation, this rule underapplies in the specific sense defined in (8). (See Section 3.4 for discussion of an example of the converse situation: a rule that does *not* make a non-vacuous change during the course of the derivation but that still does not underapply in the sense of (8).)

Examples of fed counterfeeding on focus are more commonly referred to as DUKE OF YORK DERIVATIONS (Pullum 1976; McCarthy 2003): an underlying *A* becomes *B* only to end up as *A* again. As with chain shifts, there is a comparably successful alternative to the ordering analysis of Duke of York derivations, involving the conflict-adjudication mechanism of constraint ranking in Optimality Theory (McCarthy 1999, 2003a, 2007b): in Nootka, for example, the markedness constraint driving Delabialization must be ranked higher than the markedness constraint driving Labialization. (A particular subset of Duke of York derivations is also amenable to disjunctive blocking analysis; see Section 3.2.1 below.)

3.2 Blocking

Cases of counterfeeding like those discussed above have convinced many phonologists that underapplication opacity is fully accounted for by rule ordering; after all, if a demonstrably active rule's input structural description is met by a surface representation, it makes sense to think that another, later-ordered rule created that representation. But there are also sources of underapplication other than counterfeeding, all of which have received ample attention in the phonological literature. I begin with the most obvious such source, blocking.

The very definition of blocking belies its contribution to underapplication: a rule is said to be blocked when it fails – by some principle or mechanism – to apply to a form that meets its input structural description; thus, a derivation in which a given rule $\mathbb{P}$ has been blocked may result in a surface representation to which $\mathbb{P}$ underapplies. (I say "*may* result" because another, later-ordered rule could rid the surface of representations that meet the structural description of $\mathbb{P}$. Counterfeeding can be made transparent in this way; see Section 3.4.)

I discuss here three types of blocking. The first is DISJUNCTIVE BLOCKING (Section 3.2.1), in which a rule is blocked if a strictly more specific conflicting rule is also applicable. The second is NON-DERIVED ENVIRONMENT BLOCKING (Section 3.2.2), in which a rule is blocked if its structural description is not derived phonologically or morphologically. The third is (for lack of a better term) do-something-except-when blocking (Section 3.2.3), in which a rule is blocked from creating structures that for independent reasons are not allowed to surface. (The closely-related phenomenon of do-something-only-when triggering also involves underapplication, as also noted in Section 3.2.3.) Each of these well-established phenomena has required the postulation of principles or mechanisms beyond rule ordering to account for it; given that each type of blocking (and triggering) contributes to underapplication, then, it is clear that rule ordering is insufficient to account for all cases of opacity.

3.2.1 Disjunctive Blocking Disjunctive blocking has a long and celebrated history in phonological theory (see Baković, forthcoming, for detailed discussion). It all started with the analysis of stress in Chomsky, Halle, and Lukoff (1956), Chomsky and Halle (1968), and Halle and Keyser (1971). Consider the Latin stress rules in (14), stated in standard *SPE* notation (after Anderson 1974: 97).

(14) Latin stress rules

a. $V \rightarrow [+\text{stress}]\ /\ \_\_\ C_0\breve{V}C_0^1VC_0\#$ (stress the antepenult if the penult is light)
b. $V \rightarrow [+\text{stress}]\ /\ \_\_\ C_0VC_0\#$ (stress the penult)
c. $V \rightarrow [+\text{stress}]\ /\ \_\_\ C_0\#$ (stress the ultima)

Any form fitting the structural description of one of the longer rules in (14) also fits the structural description of any shorter rule. Application of these rules to any form that meets the structural description of more than one of the rules will thus result in multiple stresses on the form, regardless of the order of the rules. However, only (14a) applies to words that fit the structural descriptions of all three rules (*pa'tricia, 'reficit*), only (14b) applies to words that fit the structural descriptions of (14b,c) but not that of (14a) (*re'fectus, re'fēcit, 'aqua, 'amō*), and only (14c) applies to words that fit its structural description and not those of the other two rules (*'mens, 'cor, 'rē*). Application of a shorter, more general rule must thus be blocked by application of a longer, more specific rule; the shorter, more general rules thus underapply, again in a way that cannot be accounted for with rule ordering alone.

Other types of examples of disjunctive blocking were identified by Anderson (1969, 1974) and Kiparsky (1973), and all such cases have since been generally accounted for by (some version of) Kiparsky's elsewhere condition (Kiparsky 1973, 1982a). (Complementary stress rules such as those in (14), on the other hand, were eventually superseded by the interaction of principles of metrical phonology, as noted by Kiparsky 1982a: 173, footnote 2.) The Elsewhere Condition imposes disjunctive ordering between two rules the structural changes of which are incompatible and the structural descriptions of which are in a proper inclusion relationship. Many, but not all, such examples can in fact also be accounted for by a Duke of York derivation (recall Section 3.1.2). An example of this kind is the interaction between Trisyllabic Shortening and *CiV*-Lengthening in English (Chomsky and Halle 1968; Kenstowicz 1994a).

(15) English rules (adapted from Kenstowicz 1994a: 218)

a. Trisyllabic Shortening e.g. *o*(ˈ*pāque*) ~ *o*(ˈ*păci*)*ty*

$$V \rightarrow \breve{V} \; / \; \begin{array}{ccc} \_\_ & C_0 & V \\ | & & | \\ ({}^{\prime}\sigma & & \sigma) \end{array}$$

b. *CiV*-Lengthening e.g. (ˈ*remĕ*)*dy* ~ *re*(ˈ*mēdi*)*al*

$$\begin{bmatrix} \text{V} \\ -\text{high} \end{bmatrix} \rightarrow \bar{V} / \begin{array}{cccc} \_\_ & C & i & V \\ | & & & | \\ ({}^{\prime}\sigma & & & \sigma) \end{array}$$

Application of these rules to forms that meet both structural descriptions results in the right surface representations, whether the rules are ordered normally (= conjunctively) or disjunctively. I explain this fact in what follows, employing as key examples the forms /*remĕdi+al*/ and /*jōvial*/ (→ |*re*(ˈ*mĕdi*)+*al*| and |(ˈ*jōvi*)*al*| after footing, respectively).

Kenstowicz (1994a: 218) advocates a disjunctive analysis, mediated by the Elsewhere Condition. The structural changes of the rules are incompatible: one rule shortens vowels while the other lengthens them. Moreover, the structural description of (*CiV*-)Lengthening is properly included in that of (Trisyllabic) Shortening: both apply to the heads of bisyllabic feet, but Lengthening applies more specifically to a [–high] head of a foot the non-head of which is an /*i*/ in hiatus. Lengthening thus blocks Shortening, and Lengthening therefore applies alone to |*re*(ˈ*mĕdi*)+*al*| (→ [*re*(ˈ*mēdi*)*al*]) and |(ˈ*jōvi*)*al*| (→ [(ˈ*jōvi*)*al*]).

Chomsky and Halle (1968: 181, 240ff.) propose a conjunctive analysis, with extrinsic ordering between the two rules.[7] Shortening applies first and gives the intermediate representations |*re*(*mĕdi*)+*al*| and |(*jŏvi*)*al*|; Lengthening then undoes the effects of Shortening in these cases, rendering the correct surface representations [*re*(*mēdi*)*al*] and [(*jōvi*)*al*]. This is a clear example of fed counterfeeding on focus (recall Nootka, Section 3.1.2, (13)): Lengthening feeds Shortening which in turn counterfeeds Lengthening, which thus underapplies.

There are other examples of disjunctive blocking that can be shoe-horned into conjunctive analyses, but only at the expense of the descriptive adequacy of the individual rules themselves. Consider, for example, the interaction between Assimilation and Deletion in Diola Fogny (Sapir 1965, Kiparsky 1973), starting with the disjunctive analysis in (16).

(16) Diola Fogny rules (disjunctive analysis, adapted from Kiparsky 1973: 98)

a. Assimilation e.g. /ni+gam+gam/ → [nigaŋgam] ‘I judge’

$$\begin{bmatrix} \text{C} \\ +\text{nasal} \end{bmatrix} \rightarrow [\alpha\text{place}] \; / \; \_\_ \begin{bmatrix} -\text{cont} \\ \alpha\text{place} \end{bmatrix}$$

b. Deletion e.g. /let+ku+ʤaw/ → [lekuʤaw] ‘they won’t go’

$$C \rightarrow \emptyset \; / \; \_\_ \; C$$

The structural description of Assimilation is properly included in that of Deletion: both apply to preconsonantal consonants, but Assimilation applies more specifically to nasals followed by non-continuants. Moreover, the structural changes of the two rules are incompatible: a consonant can either be assimilated or deleted, but not both (not discernibly, anyway). Assimilation thus applies alone when applicable, blocking Deletion.

Unlike the English rules in (15), the Diola Fogny rules as stated in (16) cannot be ordered conjunctively: under either order, Deletion will delete all preconsonantal consonants, whether or not they (were destined to) undergo Assimilation. A conjunctive analysis of the interaction between these two rules requires rules as stated and as ordered in (17).

(17) Diola Fogny rules (conjunctive analysis, adapted from Kiparsky 1973: 97)

a. Deletion′

$$\begin{bmatrix} \text{C} \\ \langle +\text{nasal} \rangle \end{bmatrix} \rightarrow \emptyset \ / \ \_\_ \begin{bmatrix} \text{C} \\ \langle +\text{cont} \rangle \end{bmatrix}$$

e.g. /na+laɲ+laɲ/ →
[nalalaɲ] ‘he returned’

b. Assimilation′

$$\text{C} \rightarrow [\alpha\text{place}] \ / \ \_\_ \begin{bmatrix} \text{C} \\ \alpha\text{place} \end{bmatrix}$$

e.g. /ku+bɔn+bɔn/ →
[kubɔmbɔn] ‘they sent’

Deletion′ deletes a nasal only if it is followed by a continuant, and otherwise deletes all preconsonantal consonants. The relevant residue of this rule – nasals followed by non-continuants – is then passed on conjunctively to Assimilation′. This means that Assimilation′ need not specify the non-continuancy of the consonant being assimilated to, because Deletion′ will have already removed the relevant strings from consideration. The continuancy of the following consonant is thus a condition on Deletion′ under this conjunctive analysis, as opposed to being a condition on Assimilation as it is in the disjunctive analysis – and herein lies the problem with the conjunctive analysis. That the following consonant must be [−cont] in order for Assimilation to apply in (16a) is a natural condition on nasal place assimilation rules (Padgett 1994), but the condition on Deletion′ in (17a) – that the following consonant should be [+cont] if the consonant-to-be-deleted is [+nasal] – is not similarly justified.

In summary, disjunctive blocking represents yet another example of underapplication that cannot be accounted for with rule ordering alone. Even factoring out examples like the Latin case in (14), instead accounting for them via the interaction of principles of metrical phonology, and examples like the English case in (15), which can be inconsequentially reanalyzed as a Duke of York derivation, there remains a residue of examples like the Diola Fogny case in (16) that are best described as involving the underapplication of a rule due to disjunctive blocking by another, rather than conjunctive ordering with respect to another.

3.2.2 Non-derived Environment Blocking A classic example of non-derived environment blocking is found in Finnish (Kiparsky 1976, 1993) and is shown in (18).

(18) Non-derived environment blocking in Finnish

		a. /tilat+i/	b. /äiti/	c. /vete/
Raising:	e → i / __ #			i
Assibilation:	t → s / __ i	≀ s	≀	s
		[tilas +i]	[äiti]	[vesi]

Glosses: (18a) 'ordered', (18b) 'mother', (18c) 'water'

The examples in (18) show that Assibilation only applies if its structural description is morphologically or phonologically derived; that is, only when the conditions for (non-vacuous) application of the rule are met by virtue of the concatenation of morphemes, as in (18a), or by the application of a prior phonological rule, as in (18c). The morpheme-final /t/ in (18a) assibilates because the conditioning vowel is in a separate morpheme; the initial /t/ does not assibilate, however – as indicated by the ad hoc '≀' symbol – because the would-be conditioning vowel is in the same morpheme. The example in (18b) has a /t/ in virtually the same phonological context as the assibilated /t/ in (18a) and yet it does not assibilate because, like the unassibilated initial /t/ of (18a), the conditioning vowel is in the same morpheme. Finally, the /t/ in (18b) assibilates because the conditioning vowel is derived by the earlier application of Raising. Assibilation clearly underapplies in Finnish, given that there are surface representations that could have undergone Assibilation but have not.

Note that the conditions that hold of non-derived environment blocking are essentially the opposite of those that hold of counterfeeding. In cases of counterfeeding, earlier-derived strings undergo a rule that later-derived strings do not; ordering this rule earlier than another rule that is responsible for those later-derived strings is thus possible. In cases of non-derived environment blocking, by contrast, later-derived strings (whether by morpheme concatenation or by phonological rule) undergo a rule that earlier-derived strings do not. Rule ordering is clearly insufficient to the task in this case: early ordering can only hope to achieve counterfeeding-type underapplication, and late ordering will if anything only increase the set of forms to which the relevant rule can apply. As the ample literature on the topic attests, some additional principle ensuring the blocking of relevant rules in non-derived environments (or, alternatively, their application only in derived environments) is necessary within rule-based serialism, in the form of either the Revised Alternation Condition (Kiparsky 1976), the Strict Cycle Condition (Kean 1974; Mascaró 1976), a combination of lexical identity rules and the Elsewhere Condition (Kiparsky 1982), or the judicious use of underspecification and feature-filling rule application (Kiparsky 1993; cf. Poser 1993). (See Burzio-Baković 2011, and references cited there for more on non-derived environment blocking and its analysis.)

3.2.3 Do-Something-Except-When Blocking Do-something-except-when blocking encompasses a wide range of cases in which a rule is blocked from creating certain structures for independently-motivated reasons. It is usually motivated by the *general absence* of a particular structure in a language, one that is otherwise expected to be created by the rule in question. It differs from disjunctive blocking in that another rule (formally related or otherwise) is generally not involved, and it differs from non-derived environment blocking in that the relevant structures are generally blocked from being created across the board, not only in non-derived environments. But it is like both of these other forms of blocking in that it involves underapplication of the blocked rule.

The earliest argument for do-something-except-when blocking was made by Kisseberth (1970). In Yawelmani Yokuts (Newman 1944; Kuroda 1967; Kisseberth 1969), short vowels are deleted between consonants *except when* such deletion would result in a tautosyllabic consonant cluster (#CC, CCC, or CC#). One way to achieve this result is, of course, to build the blocking condition into the statement of the vowel deletion rule, the environment of which can be stated as VC__CV (a "doubly open syllable"), thereby including all but those contexts in which a tautosyllabic consonant cluster is in danger of being created. Kisseberth (1970) argues that this solution misses a significant generalization uniting a suite of rules in Yawelmani phonology that are either blocked or triggered (on which see below) by the avoidance of tautosyllabic consonant clusters. He argues instead that the environment of vowel deletion could instead be simplified to C__C, with the surrounding vowels of the more complex VC__CV environment being derivative properties of a CONSPIRACY.[8] To the extent that such derivative properties can indeed be factored out of the formal statement of the environment of a conspiracy-blocked rule, then, that rule underapplies.[9]

This is also true of rules that are blocked for other do-something-except-when reasons. For example, assimilation rules are often subject to the same conditions as the underlying segment inventory itself, such that the product of assimilation cannot be a segment outside the inventory. Vowel harmony rules offer some of the most consistent evidence for this. In the vowel inventory of the Fante variety of Akan (Stewart 1967, Clements 1981, O'Keefe 2003), all vowels have a [±ATR] pair /i ~ ɪ, e ~ ɛ, u ~ ʊ, o ~ ɔ/ except the low, [–ATR] vowel /a/. As a result, the [±ATR] vowel harmony rule is blocked from applying to /a/. In this case, this blocking condition can be built in to the statement of the focus of the vowel harmony rule by stipulating that it only applies to [–low] vowels, but this has been argued since at least Kiparsky (1981) to miss a significant generalization about the relationship between conditions on harmony and conditions on the inventory. Kenstowicz and Kisseberth (1977) discussed cases like this under the rubric of the DUPLICATION PROBLEM, explaining that, as with conspiracies, the rule-based serialism model of the time was forced to view this kind of relationship as a coincidence; later work addressed the duplication problem with the STRUCTURE PRESERVATION principle (Kiparsky 1981, 1982a, 1985).

A recently proposed subclass of do-something-except-when blocking is represented by what McCarthy (2003a) calls a GRANDFATHER EFFECT, whereby a rule is blocked from creating a representation that is otherwise allowed to surface if specified underlyingly; these underlying forms are thus "grandfathered in." McCarthy uses voicing assimilation in Mekkan Arabic as an example (Abu-Mansour 1996, Bakalla 1973): underlying voiced obstruents assimilate to following voiceless ones (/ʔagsam/ → [ʔaksam] 'he swore an oath') but not vice-versa (/ʔakbar/ → [ʔakbar], *[ʔagbar] 'older'), even though voiced obstruents emerge unscathed if specified underlyingly (/ʔibnu/ → [ʔibnu] 'his son'). Grandfather effects are not independently motivated by conspiracies or inventory conditions, but McCarthy argues that they are motivated by universal markedness considerations: in the Mekkan Arabic case, the fact that voiced obstruents are marked by comparison with voiceless ones. To the extent that such markedness constraints can, like independently-motivated inventory conditions, be factored out of the rules that they block, then, rules of this kind also underapply.

Rules that are *triggered* by conspiracies or by inventory conditions (= "do something *only* when") also underapply, at least to the extent that the relevant derivative properties can be factored out of the formal statement of the environment of the triggered rule. For example, one of the conspiracy-triggered rules of Yawelmani discussed by Kisseberth (1970) epenthesizes a vowel after the first consonant of what would otherwise be a tautosyllabic consonant cluster; if the environment of the rule could thereby be reduced just to the position of epenthesis (to the effect that "in a sequence of one or more consonants, epenthesize after the first consonant"), then it would technically underapply in all sequences of one or more consonants that are not in danger of surfacing as tautosyllabic consonant clusters.

Likewise, the vowel inventory of Maasai (Tucker and Mpaayei 1955; Archangeli and Pulleyblank 1994; Baković 2000) is in all relevant respects just like the vowel inventory of Akan described above, but the unpaired low vowel /a/ only blocks leftward [±ATR] harmony; in the rightward direction, /a/ becomes [+ATR] but only by further raising and rounding to become [o]. This raising-and-rounding rule is clearly triggered by the independent absence of a [+ATR] low vowel in the vowel inventory; if the statement of the rule could thereby be reduced just to the result of rightward harmony ("raise and round a vowel that undergoes rightward harmony"), then it would technically underapply in all cases of non-low vowels.

Note that the triggered counterpart of a grandfather effect would simply be any rule the conditions for application of which can be motivated by universal markedness considerations. For example, a rule of syllable-final obstruent devoicing can be and has been argued to be motivated by the relative markedness both of voiced obstruents and of maintaining contrasts in (the rough equivalent of) syllable-final position (Lombardi 1991, 1999; Steriade 1999); this rule might thereby be reduced to the bare minimum "change (obstruent) voicing" – effectively, a rule-based imperative corresponding to a faithfulness constraint in Optimality Theory – and thus underapply when an obstruent is voiceless or not syllable-final.

Aside from the issue of underapplication, do-something-except-when blocking and do-something-only-when triggering are generally anomalous phenomena within rule-based serialism. The logic of these phenomena entails the consideration of parallel hypothetical derivations at every potential blocking or triggering turn. In order to block a rule from applying to a representation, a hypothetical application of the rule to that representation must be contemplated and found to be in violation of the blocking condition; the result is thereby discarded, and the derivation proceeds without application of the rule. In order to trigger the application of a rule to a representation, a hypothetical *non*-application of that rule must be contemplated and found to be in violation of the triggering condition; this result is thereby discarded and the derivation proceeds with application of the rule.

The necessity of these parallel derivations is rarely if ever acknowledged even in work promoting models that more explicitly acknowledge blocking and triggering (e.g. Paradis 1987; Calabrese 2005; see Odden, this volume, for discussion of some of these kinds of models). Parallel derivations are of course very much like the multiple output candidates of Optimality Theory, which was designed with blocking and triggering phenomena firmly in mind (see especially Chapters 3 and 4 of Prince and Smolensky 1993) and in which the analytical counterpart of any rule necessarily involves forced violations of some constraints; these violations roughly register the various forms of "underapplication" discussed here.

3.2.4 Summary Rules with blocking conditions underapply when they are blocked and rules with triggering conditions underapply when they are not triggered. Since satisfactory accounts of these phenomena require principles or mechanisms beyond rule ordering (the Elsewhere Condition for disjunctive blocking, one of the several proposed accounts of non-derived environment blocking, parallel hypothetical derivations for do-something-except-when blocking and do-something-only-when triggering), rule ordering is clearly insufficient to handle all examples of opacity. (This is of course true regardless of whether the additional principles or mechanisms that these phenomena require are reducible to each other or to something more general.)

3.3 *Other Examples of Underapplication*

I briefly consider here three additional examples of underapplication as defined in (8): the restriction of a rule to particular lexical classes or levels (Section 3.3.1), rule exceptions (Section 3.3.2), and rule optionality (Section 3.3.3). The identification of at least some of these types of rules as examples of underapplication is not entirely novel: rules that apply only to particular lexical classes and rules that have lexical exceptions fall into the class of "non-automatic" rules, defined by Kiparsky (1976) as those rules for which there are representations *in the immediate output of the rule* – that is, not necessarily on the surface – to which the rule could (still) apply non-vacuously.[10] The classification of at least some of these phenomena as instances of opacity may nevertheless seem counterintuitive, but as I discuss in each subsection, appropriate amendments to the definition of underapplication appear to be nothing but ad hoc. More importantly, recall that the only hypothesis tying opacity together is Kiparsky's claim that instances of it are relatively hard to learn; at a minimum, then, the relative learnability of all of these phenomena needs to be empirically determined before we begin to write some phenomenon or other in or out of the definition of opacity.

3.3.1 Restriction to Classes/Levels If a given rule applies to some but not all lexical classes or in some but not all levels, then that rule by definition underapplies with respect to the complement set of classes or levels and is thus opaque. For example, the fact that Velar Softening in English (putatively responsible for e.g. *opaque* [k] ~ *opacity* [s]) applies only to the Latinate vocabulary class means that the rule underapplies elsewhere; likewise, the fact that the rule responsible for antepenultimate main stress in English applies at Level 1 (*oríginal ~ originálity*) means that the rule underapplies at later levels (*óbvious ~ óbviousness; *obvíousness*).[11]

If this conclusion seems counterintuitive in the case of lexical classes, one could try to dismiss it by further specifying the denotation of "surface representations" in the definition of underapplication in (8) as the set of representations defined by the particular class to which the relevant rule is restricted to apply. But unless and until we can establish a relevant difference in the relative learnability of class-restricted rules and rules that underapply for other reasons (e.g. because they are counterfed), this move would be completely ad hoc.

This kind of move would not even be desirable in the case of levels because level ordering is generally an accepted mechanism for describing opaque interactions between phonological rules. For example, recall Kiparsky's (1982a) analysis of Shortening and Lengthening in English mentioned in Note 7: underapplication of Shortening is arguably due not to extrinsic within-level ordering nor to disjunctive blocking (by the Elsewhere Condition or otherwise), but rather to the independently-motivated assignments of Shortening to a cyclic level and of Lengthening to a postcyclic level. Some researchers have even claimed that *all* counterfeeding and counterbleeding interactions are due to the (independently-motivated) assignment of different rules to different levels that are serially ordered with respect to each other but within which there is no serial ordering, most notably Kiparsky (to appear) and Bermúdez-Otero (to appear); cf. McCarthy (2007b: 38ff.).

3.3.2 Exceptions If a given rule has (lexical) exceptions, then that rule by definition underapplies with respect to those exceptional forms and is thus opaque. For example, the (independently optional) rule of postnasal /t/ deletion in English (/t/ → Ø between /n/ and an unstressed vowel; see Hayes 2009: 191–192) exceptionally underapplies in the case of *intonation* for many speakers of English: [ˌĩntə̃ˈneɪʃə̃n] ~ *[ˌĩnə̃ˈneɪʃə̃n] (cf. *intellectual* [ˌĩntəˈlekʧwəl] ~ [ˌĩnəˈlekʧwəl]).

The conclusion that exceptions contribute to opacity is perhaps not so counterintuitive, but it does depend on exactly how rule exceptions are encoded in the grammar and whether the definition of underapplication opacity in (8) is sensitive to that encoding. Much as in the case of lexical class restrictions, any move to redefine underapplication to accomodate exceptions would be ad hoc unless and until a relevant difference in the relative learnability of rule exceptions and rules that underapply for other reasons is established.

3.3.3 Optionality If a given rule is optional, then by definition that rule sometimes underapplies and is thus opaque. For example, consider the optional rule of *t*/*d*-deletion in many varieties of English (see e.g. Coetzee (2004) and references therein): a form like *west* is sometimes realized as [wɛs] and other times as [wɛst]; in the latter case, *t*/*d*-deletion underapplies.

If this conclusion seems counterintuitive, one could again try to dismiss it by redefining underapplication opacity. For example, specifying the "phonological rule $\mathbb{P}$ of the form . . ." as "obligatory" would successfully, albeit stipulatively, render optional rules transparent.[12] However, this would also incorrectly exclude cases in which optional rules are uncontroversially opaque not due to their optionality but due to their interaction with other rules; see Kawahara (2002), Anttila (2006), Ettlinger (2007), and Anttila *et al.* (2008) for examples.

Another possibility is to adopt the grammar competition approach to optionality of e.g. Kroch (1989).[13] If each member of a set of possible surface realizations of a given form results from a different grammar, then optionality can be brought into the fold of transparency by saying that a rule $\mathbb{P}$ underapplies only if there are surface structures meeting $\mathbb{P}$'s input structural description *that are generated by a grammar that includes* $\mathbb{P}$. This is of course a very reasonable amendment to (or clarification of) the definition of underapplication; deliberately excluding it appears to lead to the seemingly absurd but logical conclusion that, in the case of a speaker of two languages L1 and L2, a rule $\mathbb{P}$ that is unique to L1 is opaque simply because there are surface structures meeting $\mathbb{P}$'s input structural description in L2!

There are two comments that I could make about the seeming absurdity of this logical conclusion. The first comment is that we *do* know that the grammar of one's native language can interfere with the learning of an additional language, and that at least one form of interference involves rules in the native language that do not apply in the additional language (Broselow 1983); moreover, recent research suggests that the process of acquiring multiple native languages may also involve this type of interference (Fabiano-Smith and Barlow 2010).[14] If opacity boils down to relative learnability, as Kiparsky originally suggested, then there appears to be no reason *not* to consider these types of interference between languages as types of opacity. The second comment is that, even granting the grammar competition approach to optionality, there is more than likely a continuum of conscious distinguishability between competing grammars within the same language (= less consciously distinguishable) on one end and non-competing grammars of separate languages (= more consciously distinguishable) on the other – with many points in between, of course. The relative conscious distinguishability of the grammars of separate languages vs. competing grammars within the same language could curtail the impact of opacity in the former case compared to the latter.

These comments stand apart from the by-now-familiar fact that we do not know what differences may or may not exist between the relative learnability of optionality and other forms of underapplication – and multiple language learning, for that matter. As implied throughout the preceding subsections, necessary empirical work needs to be undertaken before we jump to any conclusions about what should count as opaque and what should not.

3.4 Surface-true Counterfeeding

Another useful term introduced into the discussion of opacity by McCarthy (1999: 332) is SURFACE TRUTH: the generalization expressed by a phonological rule is *not surface true* if there are surface counterexamples to that generalization. The definition of underapplication opacity in (8) technically evaluates the surface truth of a rule, not whether the rule "applies" in all relevant derivations; however, the two notions are sufficiently co-extensive, at least in the simplest case of a pairwise interaction, that "rule $\mathbb{P}$ underapplies" and "rule $\mathbb{P}$ is not surface true" can be used interchangeably. Here I discuss an example in which a counterfed rule "underapplies" in the narrower sense that it does not apply in a relevant derivation, but in which the generalization expressed by that rule is nevertheless surface true.

In Educated Singapore English (Mohanan 1992; Anttila *et al.* 2008) there are several rules affecting word-final consonant clusters, three of which are discussed here. Epenthesis inserts a schwa between near-identical word-final consonants, much as in standard English (/reɪz+z/ → [reɪz+əz] 'raises'; cf. /bæg+z/ → [bæg+z] 'bags'). Deletion deletes a word-final plosive if it is preceded by an obstruent (/test/ → [tes] 'test'; cf. /test+iŋ/ → [test+iŋ] 'testing'). Finally, Degemination, fed by Deletion, deletes one of two word-final near-identical consonants (/list+z/ $\xrightarrow{\text{Del.}}$ |l$^{\text{D}}$is+z| $\xrightarrow{\text{Deg.}}$ [lis] 'lists').[15]

As Anttila *et al.* (2008: 185) explain, Deletion counterfeeds Epenthesis in the last of these derivations: application of Deletion results in an intermediate representation, |lisz|, to which Epenthesis is applicable, but Epenthesis does not apply; Degemination, which is also applicable, applies instead. Thus Epenthesis must apply before Deletion (= counterfeeding) and Deletion must apply before Degemination (= feeding). But despite the fact that this is counterfeeding, it does not strictly involve underapplication opacity. The fed application of Degemination ultimately removes the structural description of Epenthesis whenever Epenthesis is counterfed by Deletion, the end result being that there are in fact no surface representations to which Epenthesis could apply non-vacuously. Because Epenthesis itself is not responsible for this fact, it "underapplies" – but only in a narrower sense than justified by the definition of underapplication opacity in (8) because Epenthesis is surface true.

On the other hand, if the conspiracy behind Epenthesis and Degemination – to wit, the avoidance of surface (near-)geminates – is factored out of the formal statements of these rules in the way advocated by Kisseberth (1970), then *both* Epenthesis *and* Degemination technically underapply as defined in (8). (See Note 9 and surrounding discussion.) This is consistent with the intuition expressed by Anttila *et al.* (2008: 185) when they state that "[t]he system [of rules affecting consonant clusters in Educated Singapore English – *EB*] exhibits remarkably deep opacity," the counterfeeding interaction between Epenthesis and Deletion being one of five interactions claimed to contribute to this remarkable depth. One of the others is another counterfeeding interaction between Epenthesis and a rule of Metathesis, which amounts to exactly the same thing as the counterfeeding interaction between Epenthesis and Deletion because Metathesis also ultimately feeds Degemination. (The remaining three interactions are all examples of counterbleeding and are discussed in Section 4.4 further below.)

4 Overapplication and Counterbleeding

The definitions of overapplication opacity in (6b) and of the counterbleeding relation in (3d) are repeated (in suitably modified forms) in (19) and (20), respectively.

(19) A phonological rule ℙ of the form *A* → *B* / *C*__*D* overapplies if there are surface structures with instances of *B* derived by ℙ in environments other than *C*__*D*.

(20) 𝔹 counterbleeds 𝔸 if 𝔹 eliminates potential inputs to 𝔸 and 𝔸 precedes 𝔹.

In Section 4.1 I explain how typical examples of counterbleeding lead to overapplication as defined in (19). In Section 4.2 and Section 4.3 I discuss two types of examples of overapplication that involve (something more like) feeding than counterbleeding, and in Section 4.4 I show that counterbleeding does not always lead to overapplication as defined in (19).

4.1 *Counterbleeding*

The counterbleeding relation (20) covers situations where an earlier-ordered rule 𝔸 applies to a representation that is subsequently changed by a later-ordered rule 𝔹 such that the application of 𝔸 appears to have been unjustified; 𝔸 overapplies in such cases. Consider as an example of both counterbleeding and overapplication the following two rules of Polish.[16]

(21) Counterbleeding in Polish (Bethin 1978; Kenstowicz and Kisseberth 1979)

	a. /ʒwob/	b. /sol/	c. /gruz/
Raising:			
$\begin{bmatrix}+\text{back}\\ -\text{low}\end{bmatrix} \rightarrow [+\text{high}] \; / \; \_\_ \begin{bmatrix}+\text{voi}\\ -\text{nas}\end{bmatrix}$	u	u	
Devoicing:			
[–son] → [–voi] / __ #	p		s
	[ʒwup]	[sul]	[grus]

Glosses: (21a) ‘crib’, (21b) ‘salt’, (21c) ‘rubble’

The derivations in (21b–c) illustrate the independent action of each of the rules: Raising applies alone in (21b) and Devoicing applies alone in (21c), with no interaction in either case. In (21a), Devoicing counterbleeds Raising because the earlier application of Raising is justified in part by the fact that the following obstruent is voiced, and this critical fact about the context is subsequently changed by Devoicing. Raising thus overapplies because there are raised back round vowels that are not followed by voiced non-nasals on the surface.

The Polish case in (21) is an example of COUNTERBLEEDING ON ENVIRONMENT, because Devoicing crucially changes part of the environment that justified the prior application of Raising. There are also examples of COUNTERBLEEDING ON FOCUS, where both rules affect the same segment as in the following two rules of certain dialects of Low German.

(22) Counterbleeding in Low German (Kiparsky 1968c; Kenstowicz and Kisseberth 1971)

	a. /taːg/	b. /taːg+ə/	c. /haʊz/
Spirantization:			
$\begin{bmatrix}-\text{son}\\+\text{voi}\end{bmatrix} \rightarrow [+\text{cont}]$ / V __	ɣ	ɣ	
Devoicing:			
[–son] → [–voi] / __ #	x		s
	[taːx]	[taːɣ+ə]	[haʊs]

Glosses: (22a) ‘day’, (22b) ‘days’, (22c) ‘house’

Spirantization applies alone in (22b) and Devoicing applies alone in (22c). In (22a), Devoicing counterbleeds Spirantization because the earlier application of Spirantization is justified in part by the fact that the to-be-devoiced obstruent is voiced. Spirantization thus overapplies because there are spirantized obstruents on the surface that are not voiced.

Unlike counterfeeding, the distinction between “on focus” and “on environment” here is inconsequential; both are equally problematic or equally unproblematic for theoretical frameworks without (some analog of) serial ordering of phonological operations. Both are problematic for “classic” Optimality Theory, for example (McCarthy 1999, 2007b), and both are equally unproblematic for the Universally Determined Rule Application hypothesis of Koutsoudas *et al.* (1974), in which the rules in (21) and (22) would simply apply simultaneously to the same – in this case, the underlying – representation.

4.2 *Self-destructive Feeding*

Kiparsky (1971: 612) claims that “the unmarked status of feeding order is not subject to any serious doubt,” meaning that both of Kiparsky’s hypotheses discussed in Section 2 classify feeding as an order-to-be-diachronically-attained since it leads to both maximal utilization and transparency. But as it turns out, there exist types of feeding rule orders that involve overapplication opacity. One type is what I call SELF-DESTRUCTIVE FEEDING, in which an earlier rule feeds a later rule that in turn crucially changes the string such that the earlier rule's application is no longer justified. An example from Turkish is shown in (23).[17]

(23) Self-destructive feeding in Turkish (Kenstowicz and Kisseberth 1979)

	a. /ajag+sɯ/	b. /ʧan+s ɯ/	c. /bebeg+i/
Elision:			
s/j → ∅ / C __	∅	∅	
Deletion:			
g → ∅ / V __ V	∅		∅
	[aja + ɯ]	[ʧan+ ɯ]	[bebe +i]

Glosses: (23a) 'his foot', (23b) 'his bell', (23c) 'baby (acc.)'

Elision applies alone in (23b) and Deletion applies alone in (23c). The derivation in (23a) shows the self-destructive feeding interaction between the two: the result of Elision crucially places the stem-final /g/ in the intervocalic position that causes it to undergo Deletion (that is, Elision feeds Deletion) but the /g/ itself was a necessary part of the environment justifying the application of Elision in the first place (that is, Elision overapplies). This case is an example of SELF-DESTRUCTIVE FEEDING ON ENVIRONMENT, because Deletion crucially changes part of the environment that justified the prior application of Elision; see Baković (2007: 247ff.) for extensive discussion of an example of SELF-DESTRUCTIVE FEEDING ON FOCUS, which – somewhat counter-intuitively – does not involve overapplication.[18]

4.3 *Cross-derivational Feeding*

Another type of overapplication opacity that is not due to counterbleeding is what I call CROSS-DERIVATIONAL FEEDING. The name is meant to highlight the fact that this kind of feeding interaction cannot be handled within a single derivation; two separate derivations must be considered, one in which the feeding rule creates the conditions for the fed rule to apply in the other derivation. Because the opaque nature of cross-derivational feeding is the main thrust of Baković (2007), I attempt to merely summarize that discussion here.

Cross-derivational feeding can be demonstrated with the well-known example of the past tense alternation in English. Reviewing the facts: the past tense suffix /d/ becomes voiceless after stems ending in voiceless obstruents (e.g. /pæk+d/ → [pæk+t] 'packed') and is separated from the stem by an epenthetic vowel if the stem ends in a near-identical consonant /d/ or /t/ (e.g. /pæd+d/ → [pæd+əd] 'padded', /pæt+d/ → [pæt+əd] 'patted').

The standard analysis of this set of facts, illustrated in (24) below (see Baković 2005: 284ff. for discussion and references), has it that Epenthesis applies between word-final near-identical consonants (that is, word-final consonants that differ at most in voicing), thus applying to both /pæt+d/ (24a) and /pæd+d/ (24b). (Near-identity is loosely represented in the statement of Epenthesis with differing subscripts: $C_i \approx C_j$.) In the case of /pæk+d/ (24c), Assimilation applies to devoice the past tense suffix consonant. Given that Assimilation could in principle also have applied to /pæt+d/ (24a) if the order between Epenthesis and Assimilation were reversed (as in Educated Singapore English; recall Note 15 but see also Note 20 below), Epenthesis bleeds Assimilation in this derivation.

(24) English past tense alternation (standard bleeding analysis)

	a. /pæt+ d/	b. /pæd+ d/	c. /pæk+d/
Epenthesis:			
$\emptyset \rightarrow$ ə/ C_i __ C_j#	ə	ə	
Assimilation:			
[–son] → [αvoi] / [αvoi] __ #			t
	[pæt+əd]	[pæd+əd]	[pæk+t]

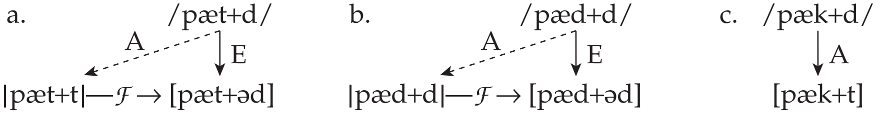
Glosses: (24a) 'patted', (24b) 'padded', (24c) 'packed'

This bleeding interaction correctly describes the fact that Epenthesis rather than Assimilation applies in (24a), but at a cost: Epenthesis must arbitrarily ignore the difference in voicing between the stem-final /t/ and the suffix /d/ – precisely the difference that would be neutralized by Assimilation were it to apply. This redundancy can be eliminated by making *strict identity* a requirement on Epenthesis (again, as in Educated Singapore English) and relying on Assimilation to provide the necessary context in (24a).

But of course Assimilation does not *actually* apply in (24a); it only *potentially* applies, but this potential appears to be sufficient to "feed" the application of Epenthesis instead. A reasonable way to model this type of interaction is with two parallel derivations, one in which Assimilation applies and another in which Epenthesis applies, as shown in (25).

(25) English past tense alternation (cross-derivational feeding analysis)

Assimilation: [–son] → [αvoi] / [αvoi] __ # Epenthesis: $\emptyset \rightarrow$ ə/ C_i __ C_i#



Assimilation is stated just as in (24) above, but Epenthesis is now stated to apply only between strictly identical word-final consonants ($C_i = C_i$). The idea here is that Epenthesis applies if and only if its structural description is met by the potential output of Assimilation; this is the case in (25a,b) – though vacuously so in (25b) – and so Epenthesis applies to those two examples. It is not the case in (25c), however, and so Assimilation applies in that example. Because the application of Epenthesis in (25a) is motivated only by the potential but not actual non-vacuous application of Assimilation, Epenthesis overapplies in this derivation in accordance with the definition of overapplication opacity in (19).

As discussed in Baković (2005, 2007), the kind of interaction illustrated in (25a) is impossible to describe with the single derivation characteristic of rule-based serialism because the potential derivation with Assimilation applying is necessary to trigger Epenthesis in the actual derivation, leading to the correct surface representation. This is in fact what makes the bleeding analysis in (24a) a necessary evil, with the arbitrary and redundant stipulation that voicing is the one feature that can be ignored in the determination of near-identity for the purposes of Epenthesis application. Cross-derivational feeding is thus yet another example of an opaque interaction that cannot be accounted for by rule ordering alone.

4.4 Mutual Bleeding

The term MUTUAL BLEEDING, following Kiparsky (1971: 600), refers to situations where a rule 𝔸 bleeds a later-ordered rule 𝔹 and where 𝔹 would also bleed 𝔸 if 𝔹 were ordered before 𝔸.[19] Whether this means that 𝔹 counterbleeds 𝔸 depends on the interpretation of the clause "𝔹 eliminates potential inputs to 𝔸" in the definition of counterbleeding in (20). The fact that 𝔸 precedes and bleeds 𝔹 in a mutual bleeding situation means that 𝔹 does not get to apply in derivations where 𝔸 applies, so there's no opportunity for 𝔹 to *actually* eliminate potential inputs to 𝔸. But if the definition is interpreted more broadly to mean that 𝔹 *in principle* eliminates potential inputs to 𝔸, then mutual bleeding counts as what we might call BLED COUNTERBLEEDING (recall 'fed counterfeeding' from Section 3.1).

Indeed, counterbleeding is often defined to more obviously encompass mutual bleeding; consider for example the following representative textbook definition.

(26) Counterbleeding (adapted from Hayes 2009: 185)

Rule 𝔹 counterbleeds rule 𝔸 when

- 𝔹 is ordered after 𝔸, and
- 𝔹 would have removed configurations to which 𝔸 applies, had 𝔹 applied first.

The "would have removed" part is the key to the inclusion of mutual bleeding, and in fact Hayes uses the following example of mutual bleeding from Lardil to illustrate counterbleeding.

(27) Mutual bleeding in Lardil (Hale 1973; Hayes 2009)

		a. /papi+ uɻ/	b. /tʲæmpæ+uɻ/
Epenthesis:	Ø → w / i __ u	w	
Elision:	V → Ø / V __		Ø
		[papi+wuɻ]	[tʲæmpæ+ ɻ]

Glosses: (27a) 'father's mother (acc. fut.)', (27b) 'mother's father (acc. fut.)'

The derivation in (27b) illustrates the independent action of Elision: the first vowel in hiatus is not an /i/, and so the second vowel is elided. In (27a), Epenthesis bleeds Elision because insertion of the glide separates the vowels in hiatus. Elision thus also counterbleeds Epenthesis here, according to the definition in (26): elision of the suffix vowel would have removed the necessary /u/ from the context of Epenthesis. A third example illustrating the "independent" action of Epenthesis is impossible to provide, given that Epenthesis applies to a proper subset of cases to which Elision is applicable. Despite its relevance in this case, note that the Elsewhere Condition (Section 3.2.1) is not needed to block Epenthesis when Elision applies because the bleeding relation between the two rules does the trick, but Koutsoudas *et al.* (1974: 8ff.) do propose that such pairs of rules are intrinsically ordered with respect to each other by the related Proper Inclusion Precedence Principle (Sanders 1974).

There is also a mutual bleeding interaction between Epenthesis and Degemination in Educated Singapore English, when the intervening Deletion rule is not involved (recall the interaction among these rules discussed in Section 3.4): /reɪz+z/ → [reɪzəz], *[reɪz]. Epenthesis clearly bleeds Degemination here by separating the members of the would-be geminate. Anttila *et al.* (2008: 185), apparently assuming the definition of counterbleeding in (26), state that Degemination also counterbleeds Epenthesis: had it applied, Degemination would have removed one of the two halves of the geminate from the context of Epenthesis.

Note that these are examples of MUTUAL BLEEDING ON ENVIRONMENT: each rule crucially disrupts the environment required for the application of the other. There are also cases of MUTUAL BLEEDING ON FOCUS, for example the following case from two different sets of dialects of German (Vennemann 1970, Kiparsky 1971: 600). In one set of dialects, the Devoicing rule already discussed in (22) bleeds a Deletion rule that deletes /g/ after nasals: /laŋg/ → [laŋk], *[laŋ] 'long (masc.)'; in the other set of dialects, the order is reversed so that Deletion bleeds Devoicing: /laŋg/ → [laŋ], *[laŋk] (cf. /laŋg+ə/ → [laŋ+ə] 'long (fem.)' in both sets of dialects, given the inapplicability of Devoicing in this case).

Mutual bleeding interactions like these obviously do not involve overapplication. Because Epenthesis bleeds Elision in Lardil, Elision does not get a chance to change the environment that justified the prior application of Epenthesis; in other words, Epenthesis in Lardil does not overapply. Likewise, because Epenthesis bleeds Degemination in Educated Singapore English, Degemination does not get a chance to change the environment that justified the prior application of Epenthesis; thus there are in fact no surface representations to which Epenthesis in Educated Singapore English has overapplied.[20] Finally, because Devoicing bleeds Deletion in some dialects of German and Deletion bleeds Devoicing in others, the bled rule does not change the environment that justified the application of the bleeding rule and so the bleeding rule does not overapply. To the extent that counterbleeding encompasses mutual bleeding, then, not all cases of counterbleeding involve overapplication.

Note that the rules involved in some examples of mutual bleeding can be implicated in a conspiracy. Epenthesis and Elision in Lardil are both hiatus-avoidance strategies, and as already noted in Section 3.4, Epenthesis and Degemination in Educated Singapore English are both (near-)geminate-avoidance strategies.[21] As discussed in Section 3.2.3, factoring out what is being avoided from the structural descriptions of the rules involved in a conspiracy inevitably results in underapplication; to the extent that mutual bleeding involves counterbleeding, then, we can conclude that some cases of counterbleeding lead to underapplication opacity.

5 Concluding Remarks

The phonology of a language is a complex system, generating a set of surface forms the ultimate token realizations of which serve as the input that language learners are exposed to and presumably use to acquire the system. To the extent that this system is composed of individual phonological rules, it is not unreasonable to assume that the easier it is to isolate the operation of those individual rules from the input, the easier it is to acquire those rules and hence the system. But phonological rules do not generally operate in isolation, nor do they tend to interact in simple pairwise ways. Although phonologists often find it useful, for expository or pedagogical purposes, to (attempt to) isolate the operation of a single phonological rule or the interaction between two rules, it is always important to be mindful of the overall system. Could the actions of other rules affect any conclusions drawn from an individual rule or interaction between rules? Could attention to other parts of the system be necessary to understand the workings of an individual rule or interaction? In the absence of solid answers to these types of questions, we have little basis beyond Kiparsky's suggestive – but by no means conclusive – diachronic evidence that it is hard to learn opaque rules; after all, such questions presumably apply not only to a phonologist's analysis of the phonology of a language but also to a learner's acquisition of one.

The resurgence of research on phonological opacity over the past 15 years or so has unfortunately not paid attention to such questions; opacity has instead been wielded as a weapon in the larger debate between proponents of rule-based serialism and proponents of alternative theoretical frameworks, Optimality Theory in particular. The debate has been sharply polarized in most respects, but there is one mistaken "fact" on which nearly all researchers on both sides (e.g. Vaux 2008, McCarthy 2007b) mysteriously appear to have decided to agree: that rule-based serialism, via its central principle of rule ordering in (2), offers a unique and unified account of opacity as originally defined by Kiparsky in (1). I have demonstrated in this chapter that this is simply not the case, unless we decide to depart from Kiparsky's agreed-upon definition of opacity and instead stipulatively (and perversely) define it as just those opaque interactions that can be described with rule ordering. Further discussions of the implications of opacity for theoretical framework comparison should either acknowledge this or provide a different, principled definition of opacity on which to base such discussions (see e.g. Bermúdez-Otero 1999, Idsardi 2000, Ettlinger 2008, and Tesar 2008, forthcoming).

This result of this demonstration is neither surprising nor a matter of concern. Kiparsky's learnability claim is really all that warrants the investigation of "opacity" as a singular notion, and there is no a priori reason to assume that the relative learnability of a phonological generalization should be reflected in the formal mechanisms used to account for its interaction with another phonological generalization that is responsible for that relative learnability, and there is even less reason to assume that any two generalizations with similar degrees of learnability should be accounted for with the same formal mechanisms. Even if there were reasons to make such assumptions, there is precious little (if any) research quantifying the relative learnability of different phonological generalizations as a function of their interactions with other phonological generalizations. In the absence of such crucial empirical work, any formal assumptions we make about opacity are bound to be tentative at best.

ACKNOWLEDGMENTS

For stimulating discussions of and helpful comments on the contents of an earlier version of this chapter, I thank the Crosslinguistic Investigations in Syntax-Phonology group at University of California–Santa Cruz (esp. Ryan Bennett and Vera Gribanova), the San Diego Phonology Interest Group at University of California–San Diego (esp. Sharon Rose), the members of Gene Buckley's seminar on opacity at Penn (esp. Josef Fruehwald), the editors of the *Blackwell Companion to Phonology* (esp. Marc van Oostendorp and Beth Hume), the editors of the *Handbook of Phonological Theory* (esp. Jason Riggle), and two anonymous reviewers. Remaining errors are mine to keep.

NOTES

1 Kiparsky (1968c) was one of the first to explicitly distinguish between these relations (see also Chafe 1968; Wang 1969; Koutsoudas *et al*. 1974), and was certainly the first to use the feeding/bleeding terminology; Newton (1971) appears to have introduced the "counter-" prefix. (Kenstowicz and Kisseberth 1971 used a "non-" prefix in the same sense; in later work, Kenstowicz and Kisseberth 1977, 1979 used "counter-.")

2 Albright and Hayes (this volume) discuss an actual example of counterfeeding and counterbleeding in different derivations, arising from the ordering of height-dependent rounding harmony before high vowel lowering in Yokuts (Newman 1944; Kuroda 1967; Kisseberth 1969; Kenstowicz and Kisseberth 1979).

3 For example, one can ask: do the definitions in (3) allow for the possibility that $\mathbb{A}$ feeds $\mathbb{B}$ because $\mathbb{A}$ bleeds some intervening rule $\mathbb{C}$ that would otherwise bleed $\mathbb{B}$? (And: is the answer intuitively correct?)

4 See also recent work by the research team associated with the Learnability Project at Indiana University (e.g. Barlow 2007; Part II of Dinnsen and Gierut 2007; Dinnsen and Farris-Trimble 2008), which documents cases of opacity that appear to arise spontaneously during the course of language acquisition. Vaux (this volume) also notes examples of spontaneous opacity arising in language games.

5 The usefulness of these terms in describing the often special phonology of reduplication was highlighted by McCarthy and Prince (1995, 1999) and was first adapted to other phenomena by Benua (1997); see Section 3.3.1.

6 The “*σ σ*” in the environment of Apocope is meant to denote the fact that the rule is blocked from creating monosyllabic words (Wilkinson 1988; Prince and Smolensky 1993), and the ad hoc feature [–apical] denotes the disjoint set of [–coronal] and [+distributed] consonants that are targeted by Deletion.

7 In Kiparsky (1982a: 154ff.), Shortening is independently classified as a cyclic rule (because it is blocked in non-derived environments; see Section 3.2.2) while Lengthening is independently classified as a postcyclic rule. Lengthening is thus *intrinsically* ordered after Shortening in this analysis; see Section 3.3.1 for more on this point.

8 See McCarthy (2002: 63) for a comprehensive bibliography of 1970s-era work on conspiracies.

9 Kiparsky (1976: 80ff.) comes to the opposite conclusion about conspiracies, stating that “the fact that languages tend to have conspiracies follows from the more general fact that languages tend to have transparent rules.” This conclusion comes on the heels of an argument against Kisseberth’s proposal that the rules participating in a conspiracy should have the function of the conspiracy factored out of their statements, Kisseberth’s claim being that this formally simplifies the grammar. Kiparsky argues that Kisseberth’s invocation of Chomsky and Halle’s (1968) formal evaluation metric is unsuccessful, but given the general lack of success of the evaluation metric – at least in the crude, feature-counting form that is relevant to the discussion – this argument does not necessarily undermine Kisseberth’s underlying proposal.

10 Optional rules are also non-automatic unless analyzed in terms of grammar competition; see Section 3.3.3.

11 Indeed, Benua (1997) adapts the terms “underapplication” and “overapplication” from Wilbur (1973) to describe just these sorts of differences in rule applicability in different levels; recall Note 5.

12 Note that the Revised Alternation Condition of Kiparsky (1976), noted briefly in Section 3.2.2, likewise stipulates that only *obligatory* neutralization rules are blocked from applying in non-derived environments.

13 I am indebted to Josef Fruehwald for raising the issues discussed in the remainder of this subsection.

14 Thanks to Cynthia Kilpatrick and Bożena Pająk for the representative references cited here.

15 In Mohanan’s analysis, Degemination only applies to clusters of *strictly identical* consonants and must thus also be fed by a voicing assimilation rule not discussed here (/list+z/ $\xrightarrow{\text{Del.}}$ |lis+z| $\xrightarrow{\text{Assim.}}$ |lis+s| $\xrightarrow{\text{Deg.}}$ [lis]). The simplification in the text does not affect the point at issue; see Section 4.3 for more relevant discussion.

16 See Buckley (2001) and Sanders (2003) for an alternative view of the Raising alternation.

17 See also Inkelas, this volume, where Paster’s (2006: 99) input subcategorization analysis of this example is summarized. (Thanks to Jorge Hankamer for instructing me on the finer points of the Deletion rule.)

18 Self-destructive feeding was first identified as an opaque feeding order in Baković (2007); the example of NON-GRATUITOUS FEEDING discussed in that article is left out here in the interests of space.

19 Thanks to Marc van Oostendorp for very helpful comments on the content of this section.

20 The mutual bleeding interaction in Educated Singapore English between Deletion and Metathesis also does not involve overapplication for the same reasons. In the end, only one of the five interactions contributing to the “remarkably deep opacity” of this system – counterbleeding between Epenthesis and Voicing Assimilation, mentioned

in Note 15 – is in fact an opaque one according to Kiparsky's definition in (6), and is the one interaction that Anttila *et al.* (2008: 194ff.) ultimately deny the factual basis of. Kenstowicz and Kisseberth (1971) put forth the idea that counterbleeding interactions between epenthesis and assimilation rules might be universally non-existent; see also Baković (2007: 245ff.) and Baković and Pająk (2008).

21 The German case is at best a less-than-clear example of a conspiracy. Both Deletion and Devoicing can do their part in ridding the surface of [ŋg] sequences, but both only actually do so in the first set of dialects described in the text; in the second set of dialects, Devoicing never gets a chance to apply to the relevant sequences. Furthermore, Devoicing more generally devoices all syllable-final obstruents, not just /g/.