Rule ordering

Preprint version. Published in M. van Oostendorp, C. Ewen, B. Hume, K. Rice, eds., *The Blackwell Companion to Phonology*. Oxford: Wiley-Blackwell, 2011, 1736-1760.

Joan Mascaró. CLT, Universitat Autònoma de Barcelona

1. The bases of rule ordering

The distributional properties of sound in natural languages are explained by resorting to a level of underlying structure in addition to the level of observed phonetic or surface representation, and to a function that maps underlying representations into surface representations. This function was conceived since the beginning of generative grammar as a (partially) ordered set of rules. A rule expresses a significant generalization about the sound structure of a given natural language. The rules of Generative phonology, as formalized in Chomsky and Halle (1968) (henceforth SPE) and subsequent work, were formalized adaptations of descriptive statements about phonology of earlier frameworks, even though their function was not the same. Both the relationship of generative rules to statements of descriptive grammars and the reasons for imposing ordering on them can be gathered from the following example taken from Halle (1962: 57-58). (1a-d) correspond to the description of Sanskrit vowel sandhi in Whitney's *Sanskrit Grammar* (Whitney 1889). (1e-h) is a formalization of the corresponding generative rules. In (1e-h) I have included only the rules that apply to front vowels for simplification.

(1) a. Two similar simple vowels, short or long, coalesce and form the corresponding long vowel. (§126)

b. An *a*-vowel combines with a following *i*-vowel to *e*; with a *u*-vowel, to *o.* (§127)

c. The *i*-vowels, the *u*-vowels and the *r* before a dissimilar vowel or a diphtong, are each converted into its own corresponding semi-vowel, *y* or *v* or *r*. (§129)

d. Of a diphtong, the final *i*- or *u*-element is changed into its corresponding semi-vowel *y* or *v*-vowe before any vowel or diphtong: thus *e* (really *ai*...) becomes *ay,* and *o* (that is *au* ...) becomes *ay*. (§131)

e. $V_iV_j \rightarrow V_i{:}$ $\quad V_i = V_j$

f. $ai \rightarrow e$

g. $i \rightarrow j$ / __ V_i $\quad V_i \neq i$

h. $i \rightarrow j$ / V_i __ V_j $\quad V_j \neq i$

The similarity of the rules to the descriptive statements is obvious. But, as Halle notices, if the ordering e-g-f shown by the arrows is imposed on the rules, "significant simplifications can be achieved." A similar comment is made in SPE (p. 18) with respect to ordering: "it [is] possible to formulate grammatical processes that would otherwise not be expressible with comparable generality." Indeed, the condition on dissimilarity of (1g) can be eliminated, since when (1g) applies all similar *VV* sequences will have coalesced by the application of (1e). Moreover, (1h) can be dispensed with because *ViV* sequences will not be turned into *eV* by (1f) since (1g) will have changed the vowel into a glide.

Since SPE relates underlying and surface reperesentations via a set of ordered rules, it follows that language variation must be due to differences in underlying representations, in the set of rules, and in their ordering. A famous example of difference in grammars stemming from different orderings of the same rules is Canadian raising, an example introduced in Halle (1962:63-64), based on data from Joos (1942), which is also discussed in Chomsky and Halle (1968: 342).

In certain Canadian and U.S. dialects the first elements in the diphtongs /□□□/, /□□□/ are raised to [□□□], [□□□] before voiceless consonants.[1] At the same time there is regular voicing of /t/ to [d] or [□] in the American English flapping environment. The interaction of these phenomena gives different results in two dialects, A and B. This causes, according to Joos, a dilemma: in a word like *writer,* which is pronounced [□□□□□□□] in dialect A, Joos's generalization that "/a/ is a lower-mid vowel ...[only] in diphtongs follwed by fortis [≈voiceless] consonants" is not true—and in Joos's view, descriptive statements are about surface representations, hence true of surface representations. Halle's solution to the dilemma stems from the recognition that statements of regularities ("rules") should be true of steps in the derivation, but need not be true of surface representations. This is the case if rules are ordered, and hence the application of a later rule can change the context that conditioned an earlier rule, as in this case, or the result of the rule itself. The statement in (2c=2b') is true of both surface representations (2d, d'), but the rule in (2b=2c') is true of (2d'), but not of (2d) which contains the sequence [□□□□], if we interpret the rule in the sense that "/□□□/, /□□□/ appear [phonetically] only before voiceless consonants." I simplify the flapping context to V__V.

(2) Dialect A Dialect B

a.		/□□□□□□□□□□□□/	a.'		/□□□□□□□□□□□□/
b.	□→□□/__[C, –voice]	□□□□□□□□□□□□	b.'□	□→□ / V__V	□□□□□□□□□□□
c.	□→□ / / V__V	□□□□□□□□□□□	c.'	□→□□/__[C, –voice]	□□□□□□□□□□□
d.	Output:	[□□□□□□□□□□□□	d.'	Output:	
[□□□□□□□□□□□□□					

Another example of grammars differing only in rule ordering is examined in Kiparsky (1982: 65-66). German devoices obstruents in coda position (3a) and simplifies □□ clusters to □□ (3b). Two of the inflective forms of the adjective meaning 'long', *lang* and *lange* contrast in two

---

[1] I transcribe the diphtong vowel as [□], and the voiced *t* as [□], following Chambers (1973, 2006); Joos's phonetic description is slightly different (basically [□] and [□], respectively. Candian raising has generated quite some discussion. Kaye (1990) casts some doubts on the existence of dialect B, that are not clearly motivated. Mielke et al. (2003) claim that the difference has been phonemicized, e.g. as v/□□□□/ vs. /□□□□/, but Idsardi (2006) argues convincingly that there are actual alternations.

dialect groups, one showing [□□□□], [□□□□], the other: [□□□], [□□□□], respectively. Application of any of the two rules renders the other rule inapplicable (mutual bleeding), therefore only the first rule applies in each odering to underlying /□□□□/:

(3) a. Devoicing [obstr] → [–voiced] / __ $\left\{\begin{array}{l}+C\\ \#\end{array}\right\}$

b. g-deletion g → Ø / [+nasal] __

c. Dialect group I

	/□□□□/	/□□□□+□/
Devoicing	□□□□□□	□—□
g-deletion	—	□□□□

d. Dialect group II

	/□□□□/	/□□□□+□/
g-deletion	□□□□	□□□□
Devoicing	—	—

Rule ordering is closely connected to rule application. As witnessed by Whitney's example, descriptive grammars and many versions of structuralist phonology implicitly assume simultaneous rule application (see Postal 1968: 140-152). This follows from the assumption that rules (or descriptive statements) are true of surface representation—are generalizations about surface representation. In simultaneous rule application the string is scanned for the structural description of each rule and all the rules whose structural description is met apply simultaneously. Chomsky and Halle (1968: 19) provide an interesting abstract example of simultaneous application, which is compared to rule ordering.[2] I adapt it with a hypothetical example. Consider rules (4a), (4b), the underlying representations (4c) and (4d), and the results of simultaneous application (4e), and ordered rules (4f, g):

(4) a. □ → □□ /— □□

b. □ → □ /—□□□□□

□

□

	Underlying	Surface		
		e□□Simultaneous application	f.□Ord. a-b	g. Ord. b-a
c.	/□□□/	□□□□□□	□□□□□	□□□□
d.	/□□□□/	t□□□□	□□□□□	□□□□□

The problem is now empirical, i.e. the question to ask is whether natural languages have input-output relations like (4c,d-e), or rather input-output relations like (4c,d-f) or (4c,d-g). The latter are cases of mutually feeding rules. With the ordering (4a)-(4b) feeding takes place in /□□□/ →□□□□□ → [□□□□]; with the ordering (4b)-(4a) feeding takes place in /t□□□/ →□□□□□ →

[2] Chomsky and Halle's (1968) example consists of the rules B → X /___Y and A → Y / ___ X and the input representations /ABY/, /BAX/.

[□□□□□]. Simultaneous application makes these feeding relationships impossible. Thus, since feeding relations are clearly observable in natural languages, rule ordering is supported, at least in front of simultaneous application.

A similar example can be constructed with mutually bleeding rules. Consider a language with palatalization of velars before *i* and backing of *i* to *u* after velars, and the underlying representation /□□/:

(5) a. □ → □□ /__ □□
b. □ → □ / k __□□□

	Underlying	Surface		
		d. Simultaneous application	e. Ord. a-b	f. Ord. b-a
c.	□□□	□□□□□	□□□□	□□

Under rule ordering, for order (5a)-(5b) we can only apply palatalization (/□□/→□□□→—). For order (5b)-(5a) we can only apply backing (/□□/→□□→—). Here simultaneous application makes bleeding impossible: both rules must apply. Simultaneous application faces another problem. A set of ordered rules assigns one and only one surface representation to any underlying representation. But consider simultaneous application of two rules, one lowering mid nasalized vowels (e.g., □□ → □□), another raising mid unstressed vowels (e.g., □ → □, if unstressed). They will force / □□□□/ → [□□□], / □□/ → [□□, under any application mode. But consider unstressed / □□□/, which satisfies both rules. Under ordering, the first rule applied always wins (we have again mutual bleeding): with the ordering lowering-raising the vowel is lowered, with the reverse ordering it is raised. Under simultaneous application, since / □□□/ meets the structural descrption of both rules, two simultaneous contradictory changes must apply to / □□□/: it has to be lowered *and* raised.

An ordering relation between rules A and B is often referred to as *extrinsic* ordering. On the other hand, *intrinsic* ordering refers to an ordering imposed by the form of the rules; when a rule is unordered with respect to another rule it can be the case that it is inapplicable unless the other rule has applied. This is typical of rewriting systems that generate an infinite set of representation from an initial symbol, as in phrase structure grammars like (6):

(6) i. S → NP VP
ii. NP → Det N
iii. VP→ V S

In (6) rule (6i) is intrinsically ordered before rule (6ii) and before rule (6iii) because their structural description cannot be met until (6i) has applied. In generative systems like (6) the ordering is generally free: different orderings generate different structures. Multiplicity of structures derived from the initial symbol (S in (5)) stem from unordering and from recursivity. On the contrary, in interpretive rule systems like classic generative phonology and other phonological theories an underlying representation is interpreted as a single surface representation (up to free variation).

Rule ordering was widely discussed in the 70's and beginning of the 80's, and a variety of different types of ordering were proposed. For details, see Iverson (1995) and especially Kenstowicz and Kisseberth (1977: 155-195).

2. Rule interaction, ordering, and applicability: feeding and bleeding

In a system of ordered rules, rules can interact: both the applicability and the result of application of a rule can depend on the application of previous rules. In 1968 Kiparsky introduced the notions of *feeding* and *bleeding* relations among rules in order to explain the direction of linguistic change. These concept have been widely usef afterwards. In this section I examine them in some detail.

Since it is not uncommon to detect some terminological inadequacies in the literature, in order to o avoid confusion I will start with some terminological observations. In Kiparsky's original terminology, *feeding* and *bleeding relations* between rules are distinguished from *feeding order* and *bleeding order*. Feeding and bleeding relations (or the terms 'X feeds/bleeds Y') are defined as functional relations between two rules, with no actual ordering between them pressuposed. *A* feeds *B* if *A* "creates representations to which *B* is applicable"; *A* bleeds *B* if *A* "removes representations to which *B* would otherwise be applicable.", where "representations" means *possible* representations (Kiparsky 1968: 37, 39). *Feeding order* and *bleeding order* are relations between rules that are in a specific order. Since feeding and bleeding *relations* are functional relations between rules, whether two rules are in *feeding* and *bleeding relation* can be determined by mere inspection of the rules.[3] I will keep this distinction (feeding/bleeding relation vs. feeding/bleeding order), but I will reserve the use of the predicates *feed* and *bleed* applied to arguments A and B for feeding/bleeding order, and I will resort to the predicates *p-feed* and *p-bleed* ('p' for 'potentially') in the case of feeding/bleeding relations. (7) icludes an illustration with our previous German example in (3):

(7)

		German, group II (ord. g-del.-Dev.)
Feeding/bleeding relation	A p-feeds/p-bleeds B	g-deletion p-bleeds Devoicing Devoicing p-bleeds g-deletion
Feeding /bleeding order	A feeds/bleeds B	g-deletion bleeds Devoicing Devoicing does not bleed g-deletion

Feeding and bleeding relations can be formally defined as follows:

---

[3] Of course one might want to to relativize these notions to a given set of representations, e.g. the lexicon. For instance, a rule *A* that centralizes the place of articulation of all consonants in word final position feeds a rule *B* that vocalizes *l* to *w* in coda position, because it can create the representation $...V\square]_{Coda}$## from . /$...V\square]_{Coda}$##/, to which *B* is applicable. But in a language with a single lateral *l*, the feeding interaction will never take place. In such cases we can say that A feeds B, but A doesn't feed B for lexicon L, or that A doesn't L-feed B. Similarly, if we relativize feeding and bleeding to specific derivations, we can say that a rule A does/does not d-feed or d-bleed A rule B meaning that the feeding or bleeding relation is not actually instantiated in that particular derivation.

(8) Feeding and bleeding relations

Rule A is in feeding relation wrt B (or A p-feeds B) iff there is a possible input I such that B cannot apply to I, A can apply to I, and B can apply to the result of applying A to I.

Rule A is in bleeding relation wrt B (or A p-bleeds B) iff there is a possible input I such that B can apply to I, A can apply to I, and B cannot apply to the result of applying A to I.

It is important to notice that in the definitions in (8) "apply" is usually interpreted as "apply non-vacuously". In the German example in (3), in dialect group I Devoicing bleeds g-deletion, (/□□□□/→/□□□□/→ —). But for the word *Bank* 'bank' whose derivation is /□□□□/→(vacuous devoicing) □□□□→—, we don't want to say that devoicng bleeds g-deletion because the input to devoicing didn't meet already its structural description. Kiparsky's (1968) terms "creates" and "removes" above already indicate that vacuous application doesn't count.

On the other hand, *feeding order* and *bleeding order* (or the terms A *feeds* B and A *bleeds* B) refer to relations between two rules A and B that pressupose both feeding/bleeding and the specific ordering A<B in the grammar. Most definitions are formulated for cases in which A inmediately precedes B, or cases in which intervening rules don't interact with A and B. In such a situation the definitions become simpler: A is in feeding/bleeding order wrt B iff A<B (i.e., A precedes B) and A feeds/bleeds B. For the general case the definitions have to be refined as follows:

(9) Feeding order and bleeding order

Let G be a grammar, A, B rules, and D a derivation of G.

a. A *is in feeding order wrt* B (or A *feeds* B) in grammar G iff
   i) A<B
   ii) There is a derivation D by G such that B would not apply to the input to A, and B applies to the output of A and would apply to all intermediate stages up to its own input.

b. A is in *bleeding order wrt* B (or A *bleeds* B) in grammar G iff
   i) A<B
   ii) There is a derivation D by G such that B would apply to the input to A, and B does not apply to the output of A and would not apply to all intermediate stages up to its own input.

When A inmediately precedes B or in cases where intermediate rules don't interact we get derivations like those in (10): (10ai) is in feeding order wrt to (10aii) because the second rule (10aii) wouldn't apply to AQ but it applies to BQ, the output of the first rule (10ai); (10bi) is in bleeding order wrt to (10bii) because the second rule (10bii) would apply to AQ but it doesn't apply to BQ, the output of the first rule (10bii).

(10) (No intervening interacting rules)

a. Feeding order			b. Bleeding order		
		AQ			AQ
i.	A → B / __Q	BQ	i.	A → B / __Q	BQ
ii.	Q → R / B__	BR	ii.	Q → R / A__	—

Feeding order (a rule A feeds a rule B) can be illustrated with the interaction of □□□□→□□□□ and Umlaut in a group of Swisss German dialects (Kiparsky 1982: 190), bleeding order (rule A bleeds rule B) with our earlier example (2a'-d'), Canadian raising in the word *writer* for dialect A:

(11) Feeding. Swiss German, Group I /□□□-li/ 'egg-diminutive'

A. □□□□→□□□ /__C, ## □□li

B. Umlaut (fronting) □□□□ (B. would not apply nonvacuously to /□□□-li/)

Bleeding. Canadian raising, dialect A

A. □→□ / V__V □□□□□□□□

B. □→□□/__[C, –voice] —— (B. would apply to /□□□□□□□/)

Consider now the cases with interacting rules intervening between A and B that motivate the definitions in (9). (12a) exemplifies feeding and (12b) bleeding cases. A and B are the rules in feeding/bleeding relation, C is the intervening rule.

(12) (Intervening interacting rules)

a. (Non-)feeding order			b. (Non-)bleeding order		
		QA			QA
A.	Q → R / __A	RA	A.	Q → R / __A	RA
C.	A → B / R__	RB	C.	A → B / R__	QA
B.	A → C / R__	—	B.	A → B / Q__	QB

In the feeding example, rule A p-feeds rule B and preceds B, but given conditions (9aii, bii), it does not feed rule B, because some rule ordered between them, namely C, undoes the change that caused the feeding (it bleeds rule B). In terms of the definitions in (9), there are representations between the two rules, in particular the input to rule B, to which the second rule cannot apply. Similarly, in the bleeding example, rule A p-bleeds rule B and would indeed bleed rule B if it were not for C which feeds rule B.

Of course a pair of rules can show non-feeding or non-bleding interactions like those in (12) in some derivations, but feeding or bleeding interactions in other derivations. I will use the terms d-feed and d-bleed when feeding and bleeding is relativized to a specific derivation. English stress provides an actual example for bleeding. Stress is assigned twice in words like *context* □□□□□□□□□□□] or *Ahab* [□□□□□□□]. But after a light syllable the second stress is removed (Arab rule, Ross 1972), as in *Arab* [□□□□□], and the destressed vowel reduces to □.

Stress bleeds vowel reduction, but in the derivation of *Arab* destressing undoes the bleeding (/□□□□/ → □□□□□□□→ □□□□□□→□ □□ □□□□□]). Here we must say that stress bleeds reduction, because there are derivations that show actual bleeding, as in □□□□□□□, but if we relativize bleeding to specific derivations, some of them do not show a bleeding interaction: in these derivations stress doesn't d-bleed reduction.[4]

There is yet another interesting case of p-feeding/bleeding with no actual feeding/bleeding. If a rule A p-feeds a rule B and precedes it, and there is a representation to which A applies and B would not apply, it is possible according to (9), to have no feeding order even if, contrary to what happens in the previous examples, B actually applies. The same is true, mutatis mutandis, of bleeding order. This happens in Duke of York derivations, which are derivations in which a rule reverses the action of a previous rule, e.g. ...A... → ...B... → ...A... Consider the derivations in (13), which contain a Duke of York (sub)derivation, highlighted in bold (notice that (12b) above is also an instance of Duke of York):

(13) (Intervening interacting rules, Duke of York derivatons)

a. Feeding order			b. Bleeding order		
		QA			**QA**
A.	Q → R / __A	**RA**	A.	Q → R / __A	**RA**
C.	A → B / R__	**RB**	C.	A → B / R__	**QA**
D.	B → A / R__	**RA**	D.	A → B / Q__	QB
B.	A → C / R__	BC	B.	A → C / Q__	—

In (13a) A p-feeds and precedes B, and B does apply, but A does not feed B, because there is an intermediate representation to which the second rule would not apply, namely RB, created by C. In fact it is the other intervening rule, D, that now feeds A. Similarly, in (13b) A p-bleeds the last rule B, but it does not bleed it, even if the rule does not apply, because of the intermediate representation QA created by C, to which the rule would apply. Here it is rule D which actually bleeds B.

Feeding and bleeding interactions have been used in different contexts and for different purposes, so it is conceivable to have slightly different changes in the definitions. One such change is desirable in cases in which usual definitions do not yield a feeding/bleding relation, and yet this relation is intuitively correct. Consider a case like (14) in which Glide Formation, Vowel Reduction, and Destressing interact in Central Catalan. (14a, b, c) show that Glide Formation affects postvocalic high unstressed vowels, but not nonhigh or stressed vowels. It also affects high unstressed vowels that are not underlying, as in (14d).

---

[4] Notice that the example is adequate only if we assume that the underlying /□□□□/ has no stress structure and the unstressed character of the second vowel is introduced by the stress rule.

(14)	a.	b.	c.	d.
	ser[□□] [□]mid	ser[□□] [□]siris	ser[□□] [□□]til	ser[□□] [□□]ciós
	'it-will-be wet'	'it-will-be Osiris'	'it-will-be useful'	'it-will-be idel'
Destressing				□□ □
Vowel Reduction				□□ □
Glide Formation	□□ □			□□ □
Output:	ser[□□] [□]mid	ser[□□] [□]siris	ser[□□] [□□]til	ser[□□] [□]ciós

Notice now that the structural description V V[+high, –stress] is met in (d) because Vowel Reduction has turned [□] into a high vowel, but also because Destressing has created the other condition for gliding. In such case we want to say that these two rules jointly feed Glide Formation. The definitions in (9) can be accordingly changed to meet such situations. Consider yet another case, where the ordering is bleeding. A representation CAD is subject to two rules C→Q / __A, D→R /A__, that are unordered with respect to each other and which precede the rule A→B / C__D. Each rule individually cannot chnage the representation CAD in such a way as to render the third rule inapplicable, but together they do. Since they are (mutually) unordered, the only way to establish a bleeding order is to say that they jointly bleed the third rule.

Notice also that a rule can both feed and bleed another rule, and stand in both feeding and bleeding order with respect to it. In Balearic Catalan stops assimilate in place to a following consonant (Place Assimilation), and the second consonant in a two-consonant coda cluster deletes before another consonant (Cluster Simplification). As shown in (15), deletion of the medial C causes bleeding when the medial C is the target of assimilation and feeding when it intervenes between the trigger and the target of assimilation.

(15) *Place Assimilation* *Cluster Simplification*

	Bleeding	*Feeding*	
Input	□□□□□□□□□□□□□	□□□□□□□□□□□□□□	'I empty trains'
	'empty trains'		
Cluster simplification	□□□□□□□□□□□□□□	□□□□□□□□□□□□□□	
Place Assimilation	——□	□□□□□□□□□□□□□□	
□ Output:	□□□□□□□□□□□□□□	□□□□□□□□□□□□□□□	

It is obvious that two rules cannot mutually d-feed each other in the same locus L: in the ordering A<B, if A feeds B, A must apply to L, and B must not be applicable; in the ordering B<A, B must apply to L, contradicting the first condition. But rules might mutually d-feed or

mutually d-bleed each other for different loci, as in our earlier examples (4) and (3), respectively. Mutual d-bleeding is possible even for the same locus, as in the example in (5), where both □ → □□ /— □□and □ → □ / k —,□when applied to /ki/ make the application of the other rule impossible.

We can now define *counterfeeding* and *counterbleeding* order, which are orderings such that the opposite order would yield a feeding and bleeding order respectively:

(16) a. A *is in counterfeeding order wrt* B (A *counterfeeds* B) in grammar G iff
i) A<B
ii) B is in feeding order wrt A
a. A *is in counterbleeding order wrt* B (A *counterbleeds* B) in grammar G iff
i) A<B
ii) B is in bleeding order wrt A

Counterfeeding order can be illustrated with Swiss German (group II the dialect with the ordering opposite to the one shown in (11)), counterbleeding order with Canadian raising, dialect A (2a):

(17) *Counterfeeding*

Swiss German, Group II	/□□□-li/	'egg-diminutive'
A. Umlaut (fronting)	——	
B. □□□□→□□□ /__C, ##	□□ □□□	(B. would not apply to /□□□-li/)

Opposite, feeding ordering would yield: /□□□-li/→□□□-li →□□□-li

Counterbleeding

Canadian raising in dialect A

	/□□□□□□□/	
A. □→□□/__[C, +tense]	□□□□□□□	
B. □→□ / V__V	□□□□□□	(B. would apply to /□□□□□□□/)

Opposite, bleeding ordering would yield: /□□□□□□□/→□□□□□□□□→□—

It should be observed that the fact that two rules A, B do do not have a feeding or bleeding interaction does not mean that they don't interact. In (18) rule A deletes final consonants, while B stresses a last heavy syllable, otherwise the penult. different orderings give different results, but the interaction isn't either a feeding or a bleeding relation.

(18)	A. C → Ø / __##	/satopek/ —A→ satope —B→ satópe
	B. V →[+stress] / __C_0(VC_1)	/satopek/ —B→ satopék —A→ satopé

"Counter" orderings have important properties. Assume the simple case where rules A and B are adjacent, and A<B. Since in feeding order there must be at least one input I such that A is applicable to I, B is not, and B is applicable to the output of A (9), it follows that in the corresponding counterfeeding order where B<A there must be an input (namely I) to which the first rule, now B, does not apply and to which the second rule, now A, applies. Hence the generalization expressed by B does not appear in the output: we can say, using McCarthy's (1999) terms that it is *not surface-true,* it is not true of the of the output of A, usually the surface representation. In bleeding order there must be by definition at least one input I such that both A and B are applicable to I, and B is not applicable to the output of A; it follows that in the corresponding counterfeeding order with B<A there can be an input to which the first rule, now B, applies and to which the second rule, now A, does not apply. Hence the generalization expressed by A about the input I does not appear in the output: following McCarthy we can say that it is *not surface-apparent,* because the generalization A about I is not apparent in the output o A, usually the surface representation.

(19)	Feeding		Counterfeeding		Bleeding		Counterbleeding	
		I		I		I		I
	B	I'	B	—	B	I'	B	I'
	A	I''	A	I'	A	—	A	—, I''

It is important to notice that the existential quantification in the definitions in (9) of feeding and bleeding orders (hence also of counterfeeding and counterbleeding orders) allows for the existence of multiple feeding and bleeding relations between two rules. For feeding, and given two ordered rules A<B, the requirement (9aii) that there be an input I whose derivation D meets the conditions required in (9aii) does not prevent the existence of another input I' that meets the condition (9bii) for bleeding. Hence A can both feed and bleed B (and B can both counterfeed and conterbleed A). Similarly, when we consider opposite orderings it can be the case that A feeds B and B feeds A (mutual feeding), and that A bleeds B and B bleeds A (mutual bleeding), or that A feeds B and B bleeds A. The remaining logical possibilities are that a rule feeds or bleeds another rule but doesn't bleed or feed it, respectively, and that there is no interaction in the opposite order.

The only restrictions in these cases of multiple interaction regards orders relative to a given single representation R. Obviously, a rule A cannot both feed and bleed a rule B for the same input I in the same locus of application. Less obviously, if A feeds B in the derivation of of I, then B cannot feed or bleed A in the corresponding derivation of I. The only possibility is that A bleeds B and B bleeds A. The reason, for the first case, is that if A feeds B in the derivation of I, by definition B is not applicable to I, and if at the same time B feeds or bleeds A, then also by definition B should be applicable to I—hence a contradiction. But mutual S-bleeding is possible, as witnessedd by our German example in (3): Devoicing bleeds g-deletion for input I=/□□□□/ (/□□□□/ → /□□□□/→ —), and g-deletion bleeds Devoicing in the same input I=/□□□□/ (/□□□□/ → /□□□/ → —).

Applicability of rules is also related to disjunctive ordering and to the Elswhere Condition. SPE distinguishes between conjunctive ordering of two rules (the normal mode of application) and disjunctive ordering, where only one of the rules can apply. In SPE conjunctive ordering and disjunctive ordering derive from abbreviatory devices: ...{A B}... corresponds to two conjunctively ordered rules, ...A... and ... B..., ...A(B)... corresponds to two disjunctively ordered rules, ...B... and ...AB...

Disjunctive ordering was later derived as an effect of a more general principle, the Elsewhere Condition, originally formulated by Kiparsky (1973: 94). This is the revised formulation in Kiparsky (1982:136):

(20) Rules A, B apply disjunctively to a form Φ 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 Φ is distinct from the result of applying B to Φ.

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

Consider, as an illustration, the interaction of lengthening and shortening in English (Myers 1987, Halle 1995). Lengthening and shortening affect stressed nonhigh vowel in branching feet. They are lengthened if, in addition, they are followed by an open syllable ending in an /i/ which in turn must be followed by a vowel in hiatus (21c,d; lengthened vowel in boldface). Otherwise, they are shortened (21e):

(21)	a. V[–hi□→□□□ / [__ C_0 i $]_{StrF}$ V□□	c.	re[m**e**di]al	[r**a**di]al
			co[l**o**ni]al	De[v**o**ni]an
	b. V[–hi□→□□ / [__ C_0 V $]_{StrF}$	d.	[trivi]al	di[vini]ty
			[natu]ral	ex[plana]tory

(21a, b) are two rules applying in the general context "disyllabic stress foot with stressed nonhigh vowel." (21a) being more restrictive, it applies only when in addition the second vowel is *i* and it is in hiatus with a following vowel. Since [V[–hi□ C_0 i $]_{StrF}$ V is a subset of [V[–hi□□ C_0 V $]_{StrF}$ (21a) and (21b) are in an elswhere relationship: (21a) is the specific rule applying in [V[–hi□ C_0 i $]_{StrF}$ V and (21b), the general rule, applies elswhere in the general common context.

Kiparsky's (1982) use of the Elsewhere Condition to predict blocking of general morphological processes by more specific ones and strict-cyclic effects. The plural rule in (22) applies to the morphlogically pluralized form person_{Pl}, (← person), but it shouldn't apply to the lexical plural people_{Pl}. Kiparsky assumes tha every lexical item L is an identity rule L ↔ L. Since [people] $_{Noun, +Pl}$ is a subset of the set defined by __] $_{Noun, +Pl}$, the identity rule applies (vacuously) and blocks the application of the plural rule. [person] $_{Noun, +Pl}$, on the other hand, is not a lexical item but the pluralization of the lexical item [person] $_{Noun}$, and the plural rule applies regularly.

(22)		a. [people] $_{Noun, +Pl}$	[person] $_{Noun, +Pl}$
b.	[people] $_{Noun, +Pl}$ ↔ [people] $_{Noun, +Pl}$	[people] $_{Noun, +Pl}$	
	[person] ↔ [person]		——
c.	Ø → z / __] $_{Noun, +Pl}$	——	[person] $_{Noun, +Pl}$ -z

The Strict Cycle Condition prevents rule application in non-derived environments. Thus the shortening rule (cf. (21b)) shortens [[s□n]ity] to [[s□□n]ity] because it is derived, but fails to apply to [nightingale] (*[n□□ghtingale]) because it is underived. Again the identity rule [s□n] ↔ [s□n] cannot block TSS because the structural description of the identity rule [s□n] and the structural description of shortening, [V C_0 V $]_{StrF}$, are not in proper inclusion relation, but in intersection. But [nightingale] properly includes [V C_0 V $]_{StrF}$. Cyclic effects in lter cycles/levels derive from the assumption that the output of every cycle/level is a lexical item, hence an identity rule.

Given OT, some properties of Elswhere-type interactions follow as a theorem, Panini's Theorem on Contraint-ranking (PTC) (Prince and Smolensky 1993/2004: §5.3, 7.2.1). PTC relates the activity of two constraints, S and G, in a constraint hierarchy relative to an input *i*. Assume that S applies nonvacuously to *i* (i.e., it distinguishes the set of candidates Gen(*i*)). If G>>S and G is active on *i* (i.e., it distinguishes the set of candidates Gen(*i*) when it apples) then S is not active on i.[5]

3. Serial and parallel approaches

Rule interactions of the sort just discussed have become important in the theoretical comparative analysis of serial and parallel approaches, in particular in relation to opacity [cross-reference to EB's chapter]. If we compare a standard serial theory like SPE and OT, pure feeding and pure bleeding order effects (i.e., those that are not also counterfeeding or counterbleeding) can be derived from both. Consider the well-known case of interaction of of e-raising and t→s in Finnish (Kiparsky 1973: 166-172):

(23)		□□□□	'water-nom.sg.'	□□□□□-□	'wanted'
a.	e → i / __##	□□□□		—	
b.	t → s / __ i	□□□□		□□□□□-□	

Because both (23a) and (23b) are statements that are true of surface forms, constraints of the from *e##, *□□, dominating conflicting faithfulness constraints, together with other constraints determining the choice of □□ □□□□□, will derive the output of /□□□□/, /□□□□□-□/. But consider now the example of mutually feeding rules in (4). Since they are mutually feeding they are also counterfeeding in both orderings:

(24)				/□□□/□□	/□□□□/	
	a.	□ → □□ / — □□	c.	□□□□□□	□—	
	b.	□ → □ /—□□□□	d.	□□□□□□	□□□□□	□□, not □□□!

In order to derive □□□□ from□□□□□□□*□□□ must dominate FAITH *e*, and *□□ must dominate FAITH *t* (25a). But this ordering makes □□□□ from□□□□□□ impossible to obtain, because high-ranked *□□ will force the victory of □□□□□ as shown in (25b).

[5] For a clear and interesting discussion of the relation between the Elswhere Condition and PTC, see Bakovic (2006); see also Prince (1995).

(25) a.

/□□□/	*□□□	*□□	Faith e	Faith t
□□□		*		
☞□ □□□□			*	
□□□□	*		*	*

b.

/□□□□/	*□□□	*□□	Faith e	Faith t
□□□□	*			
√□ □□□□		*	*	
☞□ □□□□□			*	*

Therefore in counterfeeding ordering the first rule is not "surface true": its structural description □□ appears in the phonetic representation without the corresponding structural change (□ → □□), because the second rule has introduced changes that would allow the application of the first rule. When the ordering is counterbleeding, the generalization expressed by the first rule is not "surface apparent" In Canadian raising, dialect A, the change □□□t →□□□t does not appear as such in the phonetic representation of *writer* because the second rule has changed the result of the change, turning the triggering voiceless □ into □:

(26) a. /□□□□□/ /□□□□□□□/□

b. □□□□→□□□□/__[C, –voice] □□□□□□ □□□□□□□/

c. □→□ / / V__V □□□□□□ □□□□□□□□ □□□, not □□□□!

Here in order to obtain □□□□□ in *type,* both *□□□[C, –voice] and *V□V must be active. But for *writer* the input /□□□□□□□/ where both constraints are relevant cannot have as output □□□□□□□□□ because the candidate □□□□□□□□also satisfies both markedness constraints and is, in addition, more faithful to the input:

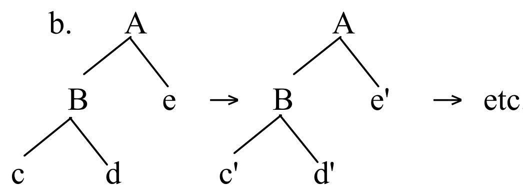
(27)

/□□□□□□□/	*□□□□[C, –voice]	*V□V	Faith □□□	Faith t
□□□□□□□	*	*		
√ □ □□□□□□ □		*	*	
☞ □ □□□□□□ □				*
□□□□□□□□			*	*

4. Cyclic ordering

In cyclic application the input to the set of rules is a phnological representation organized by constitutent structure, which is represented in (28a) by (proper) bracketing, and by a tree in (28b); A, B are categories, and d, e, c are phonological strings, matrices, or structures.

(28) a. $[_A[_B\ c\ d\ ]_B\ e]_A$ b.



The set of (partially) ordered rules $\mathfrak{R}$ applies first separately to the innermost constituents, i.e. to c, to d, and to e, giving as a result c', d', e'. This is the first cycle. We now proceed to the second cycle, the next degree of embedding, namely B, and apply the set of rules to its domain, the concatenation c'□□□□□d', whose output id (c'□□□□□□')'. The domain of the following (and in this case the final) cycle is whatever is dominated by A, namely (c' □□ d')' □ e.

The cycle was first proposed in phonology (Chomsky, Halle, and Lukoff 1956) to deal with stress in compounds like *black board eraser* showing primary-tertiary-secondary stress distinctions:

(29) $[_A[_B\ \overset{1}{[\text{black}]}[\ \overset{3}{\text{board}}]]_B\ \overset{2}{[\text{eraser}]}]_A$

After stress has applied to individual words, the compound stress rule locates stressed vowels and maintains primary stress on the leftmost stressed vowel and weakens other stresses by one degree. After assigning vacuously primary stress **1** in the first cycle to [bla$_1$ck], to [boa$_1$rd], and to [era$_1$ser], it applies in the second cycle in the domain B to [bla$_1$ck boa$_1$rd], yielding [bla$_1$ck boa$_2$rd], and in the last cycle in the domain A to [bla$_1$ck boa$_2$rd era$_1$ser] to give the final [bla$_1$ck boa$_3$rd era$_2$ser].

Ciclicity was later applied in syntax as a result of the elimination of generalized transformations and the generation of embedded sentences by base rules (Chomsky 1965). Later on, Chomsky (1973) proposed a limitation on cyclic application in syntax, the Strict Cycle Condition (SCC, or "strict cyclicity") by which no rule can apply to a constituent I in such a way as to affect solely a subconstituent of I. Kean (1974) presented two cases that argued for the application of the SCC also in phonology. In (28), for instance, in the second cycle, cycle B, a rule cannot apply to the domain of B if it just affects c'. An actual example is the interaction of Glide Formation and Destressing in Catalan. Glide Formation applies to postvocalic high vowels. In *produirà* 'it will produce' it cannot apply to [$_2$[$_1$□□□□□□□]□□□] at cycle 1 because postvocalic □□ is stressed. In cycle 2 a following stress causes destressing of □□. Therefore at cycle 3, and at later cycles, the sequence □□□meets the structural description of the rule; but □□ is enterely within cycle 1 and the SCC blocks application, resulting in [□□□□□□□□□], *[□□□□□□□□□].

The SCC was further refined in Mascaró (1976:1-3) as in (27). Case (30Ba) corresponds roughfly to the SCC as formulated in Chomsky (1973) and used by Kean (1974).

(30) Given a bracketed expression $[_n \ldots [_{n-1} \ldots, [_1 \ldots]_1, \ldots]_n \ldots]_n$, and a (partially ordered) set of cyclic rules C,

A. C applies to the domain $[_j \ldots]_j$ after having applied to the domain $[_{j-1} \ldots]_{j-1}$, each rule in C applying in the given order whenever it applies properly in j.

B. *Proper application of rules.* For a cyclic rule R to apply properly in any given cycle j, it must make specific use of information proper to (i.e. introduced bu virtue of) cycle j. [This situation obtains] if either a, bor c is met:

a. R makes specific use of information uniquely in cycle j. That is, it refers specifically to some A in $[_j$ XAY $[_{j-1} \ldots]$Z] or $[_j$ Z $[_{j-1} \ldots]$ XAY].

b. R makes specific use of information within different constituents of the previous which cannot be referred to simultaneously until cycle j. R refers thus to some A, B in $[_j$ X $[_{j-1} \ldots$A...] Y $[_{j-1} \ldots$B...] Z].

c. R makes use of information assigned on cycle j by a rule applying before R.

A states the general procedure for cyclic application; B gives the conditions for proper application: morphologically derived environments in inflection (30Ba), derived environment by compounding or syntax (30Bb), and rule derived environments (30Bc). Effects of derived environments on application of processes, irrespectively of the theoretical mechanism they derive from, are usually referred to as *derived environment effects* (DEE). We just saw a case under (30Ba). The rule of t→s assibilation in Finnish illustrates both (30Ba) and (30Bc). Assibilation applies in morphologically derived environments (31c): the structural description □□ is met by material in the root cycle and in the inflected word cycle. It also applies in rule-derived environments (31d): here the structural description □□□ is met because at its cycle of application a rule has created it. But it fails to apply in the nondervided environments (31e), because none of the conditions for proper application in (31B) is met:

(31) a. t → s / __ i

b. e → i / __##

c.	halut-a	'to want'	halus-i	'wanted'
			turpot-i	'swelled'
			hakkat-i	'hewed'
d.	vete-nä	'water-essive.sg.'	vesi (← veti ← vete)	'water-nom.sg.'
	käte -nä	'hand-essive.sg.'	käsi (← vkäte← käte)	'hand-nom.sg.'
e.	tila	'place, room'		

äiti 'mother'
itikka 'mosquito'

An instance of (30Bb) is the application of Glide Formation in Central Catalan to vowels of different words. As we just saw, Glide Formation applying to postvocalic high vowels is blocked in *produirà,* [$_2$[$_1$□□□□□□□]□□□] . Consider now *produirà oxidació* 'it will produce oxidation':

(32)	*Input to:*	[$_4$[$_2$[$_1$□□□□□□□□]□□□]□[$_3$ [$_2$[$_1$□□□□□□□]□□]□□□□]]□
	Cycle 2	[$_2$[$_1$□□□□□□□□]□□□] - [$_2$[$_1$□□□□□□□]□□]□□□□] □
	Cycle 3	[$_2$[$_1$□□□□□□□]□□□] - [$_3$ [$_2$[$_1$□□□□□□] □□□□□□□]]□
	Cycle 4	[$_4$[$_2$[$_1$□□□□□□□□□□]□[$_3$ [$_2$[$_1$□□□□□□]□□]□□□□]]

In the second word, at cycle 2 the initial □□ is destressed by a following stress and becomes □ by a rule of vowel reduction. At cycle 4, the sequence □□ meets the structural description of glide formation and the SCC does not block Glide Formation, the application being proper by (30Bb), because □□ is not within the domain of a single previous cycle: □□is in cycle 2, □ is in cycle 3. Hence the rule applies yielding □□.

The SCC was supposed to apply to cyclic, obligatory neutralization rules and covered DEEs. These were previously covered by the Alternation Condition proposed by Kiparsky (1974:65), according to which "neutralization processes apply only to derived forms ... [ie.] if the input involves crucially a sequence that arises in morpheme combinations or through the application of phonological processes." Cyclic application and derived environment effects were reformulated within Lexical Phonology through lexical strata and postlexical phonology, which correspond to cycles, and to the effect of the Elswhere Condition from which DEE are derived. In Stratal Optimality Theory [cross-reference to RBO's chapter], cycles correspond to strata to which Gen and Eval apply succesively. Within standard OT, output to output faithfulness constraints (Benua 1997) ensure similarity of larger constitutents to its inner components. Strict cycle effects (DEEs) are also obtained by local conjunction of markedness and faithfulness constraints (Lubowicz 1999). To see how DEEs are derived from local conjunction, consider the interaction of Velar palatalization and Spirantization in Polish (Lubowicz 1999: §3). We get the following descriptive generalizations. Spirantization applies to rule-derived □□ (33a), but not to underlying □□ (33b); similarly, in (33c) Velar palatalization applies only to morphologically-derived velar+e,i sequences.

(33)	a.	/rog-ek/ 'horn'	b. /ban□□-o/ 'banjo'	c. /xemik-ek/ 'chemist-dim.'
	Velar palatalization	ro□□-ek	——	xemi□□-ek *blocked in* xe
	Spirantization	ro□□-ek	*blocked*	——
	Output:	ro□□-ek	ban□□-o	xemi□□-ek
			*ban□□-o	* □□emi□□-ek

Let us examine rule-derived environments first. Given the ranking *□□ >> IDENT(cont) we will normally have the mapping □□ → □□. The difference from derived vs. non-derived environments stems from the fact that in the first case the mapping is /□/□→□□ → □□, whereas in the second case it would be /□□□/□→ □□. The candidate with □□ deriving from□□□/□ will violate both *□□ and IDENT(cor), hence the constraint conjunction *□□ & IDENT(cor)□□But if □□ is underlying, *□□ will be violated, but not and IDENT(cor), therefore the conjunction *□□ & IDENT(cor) will be satisfied.

For morphologically derived environments, as in the example in (33c), Lubowicz uses conjunction of markedness and ANCHOR. Velar palatalization applies to the morphologically-derived sequence □-□, but not to the non-derived sequence □□ in □□□□□-□□. Since the velar □ is stem-final but not syllable-final in [□□.□i.□]$_{Stem}$-□□, the sequence □]$_{Stem}$-i will violate R-ANCHOR(Stem; σ), and it will also violate PAL, the constraint against velar+{e,i} sequences that forces palatalization. It will therefore also violate [PAL & R-ANCHOR(Stem; σ)]$_D$. But since morphologically-underived xe satisfies R-ANCHOR(Stem; σ), the conjunction will be satisfied in this case and palatalization will not take place.

References

Bakovic, Eric. 2006. "Elsewhere effects in Optimality Theory." In Eric Bakovic, Junko Ito, and John McCarthy, ed. *Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince,* Santa Cruz: UC Santa Cruz Department of Linguistics, 23–70.

Research Center. Http://repositories.cdlib.org/lrc/prince/4.Benua, L. (1997). *Transderivational identity: phonological relations between words*. PhD dissertation, University of Massachusetts, Amherst. Published New York: Garland 2000.

Chambers, J. K. 1973. Canadian Raising. *The Canadian Journal of Linguistics / La revue canadienne de linguistique,* 18: 113-35. Reprinted in Chambers (1975: 83-100).

Chambers, J. K. 1975. *Canadian English: Origins and Structures.* Toronto: Methuen.

Chambers, J.K. 2006, "Canadian Raising retrospect and prospect." In Peter Avery, Alexandra D'Arcy, Elaine Gold and Keren Rice, eds., *Canadian English in the global context*, special issue of *The Canadian Journal of Linguistics / La revue canadienne de linguistique,*, 105-118.

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

Chomsky, Noam, Morris Halle & Fred Lukoff 1956. "On Accent and Juncture in English." In Morris Halle, Horace Lunt, Hugh McLean and Cornelis van Schooneveld, eds., *For Roman Jakobson. Essays on the occasion of his sixtieth birthday*, The Hague: Mouton, 65-80.

Chomsky, Noam. 1965. Aspets of the Thoery of Syntax. Cambridge: MIT Press.

Chomsky, Noam. 1973 " Conditions on Transformations." In Steve R. Anderson and Paul Kiparsky, eds., *A Festschrift for Morris Halle*, New York: Holt, Rinehart & Winston.

Halle, Morris (1962). "Phonology in generative grammar." *Word* 18: 54- 72. Whitney, William Dwight

Halle, M. 1995. "Comments on Luigi Burzio's 'The rise of optimality theory'." *Glot International* 1, 9/10.

Idsardi, William. 2006. "Canadian Raising, opacity, and rephonemicization." *The Canadian Journal of Linguistics / La revue canadienne de linguistique,* 51(2/3), 119-126.

Iverson, Gregory K. 1995. "Rule orderinng." In John Goldsmith, ed., *The Handbook of Phonological Theory*. Oxford: Blackwell, 609-614.

Joos, Martin (1942). "A phonological dilemma in Canadian English." *Language* 18: 141-144. Reprinted in Chambers (1975: 79–82).

Kaye, Jonathan. 1990. "What ever happened to dialect B?" In Joan Mascaro and Marina Nespor, eds., *Grammar in Progress: Glow Essays for Henk van Riemsdijk*, Foris: Dordrecht, 259-63.

Kean, Marie-Louise (1974). "The Strict Cycle in phonology." *Linguistic Inquiry* 5:179-203.

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

Kiparsky, Paul. 1973. "Abstractness, opacity, and global rules." Distributed by Indiana University Linguistics Club.

Kiparsky, Paul. 1973. "Elswhere in Phonology." In Steve R. Anderson and Paul Kiparsky, eds., *A Festschrift for Morris Halle*, New York: Holt, Rinehart & Winston.

Kiparsky, Paul. 1968 "Linguistic universal and linguistic change." In Emmon Bach and Robert T. Harms (eds.), Universals in Linguistic Theory , New York: Holt, Rinehart and Winston, 170-202. Reproduced in Kiparsky (1982: 13-43).

Kiparsky, Paul. 1973. "How Abstract is Phonology." In O. Fujimura, ed., *Three Dimensions of Linguistic Theory*, TEC Publishing Co., Tokyo, 5-56. Reproduced in Kiparsky (1982: 119-163).

Kiparsky, Paul. 1974. "On the evaluation measure." *Papers from the Parasession on Natural Phonology*. Chicago: Chicago Linguistics Society. 328-37. Reproduced in Kiparsky, Paul. (1982: 189-215).

Kiparsky, P. 1982a. “Lexical Phonology and Morphology.” In The Linguistics Society of Korea, ed., *Linguistics in the Morning Calm*, Seoul: Hanshin Publishing Co., pp. 3-91.

Kiparsky, Paul. 1982b. *Explanation in Phonology* In Linguistics in the Morning Calm,. Dordrecht: Foris Publications.

Lubowicz, Anna, 2002. "Derived environment effects in Optimality Theory." *Lingua* 112, 243-280.

Mascaró, Joan (1976) Catalan Phonology and the Phonological Cycle. Ph.D. dissertation, MIT. Indiana University Linguistics Club, 1978.

McCarthy, John. 1999. "Sympathy and phonological opacity." *Phonology* 16: 331–399.

Mielke, Jeff, Mike Armstrong and Elizabeth Hume. 2003. "Looking through opacity." *Theoretical Linguistics* 29.1-2: 123-139.

Myers, S. 1987. "Vowel shortening in English." *Natural Language and Linguistic Theory* **5**, 485–518.

Postal, Paul. 1968. *Aspects of Phonological Theory*. New York: Harper & Row.

Prince, Alan. 1997. "Elsewhere and Otherwise." *Glot International* 2:1, 23-24. ROA-217.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar*. Rutgers University Center for Cognitive Science: Technical Report RuCCS-TR-2. ROA-537. Published 2004, Malden, MA: Blackwell.

Ross, John R. 1972. "A reanalysis of English word stress (part I)." In Michael K. Brame, ed., *Contributions to generative phonology*. Austin: University of Texas Press, 229-323.

Whitney, William Dwight (1889). *Sanskrit grammar, including both the classical language, and the older dialects of Deda and Brahmana.* Reprinted by Harvard University Press: Cambridge, Mass. 1975.