Bibliographic Details

The Blackwell Companion to Phonology

Edited by: Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice

eISBN: 9781405184236

Print publication date: 2011

74. Rule Ordering

Joan Mascaró

Sections

1 The bases of rule ordering

The distributional properties of sound in natural languages are explained by appeal to a level of underlying structure in addition to the level of observed phonetic or surface representation (CHAPTER 1: UNDERLYING REPRESENTATIONS), and to a function that maps underlying representations into surface representations. This function has been conceived since the beginning of generative grammar as an ordered set of rules. In this chapter I will first introduce the main properties of rule ordering and the arguments for ordering rules (§1), and I will review various proposals to modify rule ordering in early generative phonology (§2), including cyclic ordering (§3). In §4 I discuss feeding, bleeding, and similar interactions in more detail, §5 discusses serial ordering and parallel approaches, and §6 draws some conclusions.

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**; *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)–(1d) are taken from the description of Sanskrit vowel sandhi in **Whitney (1889)**. The rules in (1e)–(1h) are a formalization of the corresponding generative rules. For simplification, in (1e)–(1h) I have included only the rules that apply to front vowels.

(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 diphthong are each converted into its own corresponding semi-vowel, *j* or *v* or *r*. (§129)
d. Of a diphthong, the final *i*- or *u*-element is changed into its corresponding semi-vowel *j* or *v* before any vowel or diphthong: thus *e* (really *ai* . . .) becomes *aj*, and *o* (that is *au* . . .) becomes *ay*. (§131)
e. $V_iV_j \rightarrow V_i\text{:}$ $\quad V_i = V_j$
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) is imposed on the rules in (1), "significant simplifications can be achieved." A similar comment is made by **Chomsky and Halle (1968**: 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 V*i*V sequences will not be turned into *e*V by (1f), since (1g) will have changed the vowel into a glide. We find here one of the main reasons for imposing ordered rules: ordering allows for simplification of grammars and for a better expression of linguistically significant generalizations. Another typical argument in favor of rule ordering is language variation. Since *SPE* relates underlying and surface representations 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).[1]

In certain Canadian and US dialects the first elements in the diphthongs /aɪ aʊ/ are raised to [ʌɪ ʌʊ] before voiceless consonants.[2] At the same time there is regular change of /t/ to [ɾ] 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 [rʌɪɾə] in dialect A, Joos's generalization that "/a/ is a lower-mid vowel … [only] in diphthongs followed 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. In other words, rule ordering solves Joos's dilemma. (2) shows the derivation of *typewriter* with the diphthong /Aɪ/ both before non-flapping (/p/) and flapping (/t/) environments in both dialects. The statement in (2c) (also in (2f)) is true of surface representations (2d) and (2h), but the rule in (2b) (also in (2g)) is true of (2h), but not of (2d), which contains the sequence [ʌɪd], if we interpret the rule in the sense that "/ʌɪ ʌɪ/ appear [phonetically] only before voiceless consonants." In (2) I simplify the flapping context to V___V.

(2) *Dialect A*

a.		/taɪpraɪtə/
b.	a → ʌ / __ [C, –voice]	tʌɪprʌɪtə
c.	t → ɾ / V __ V	tʌɪprʌɪɾə
d.	*output*	[tʌɪprʌɪɾə]

Dialect B

e.		/taɪpraɪtə/
f.	t → ɾ / V __ V	taɪpraɪɾə
g.	a → ʌ / __ [C, –voice]	tʌɪpraɪɾə
h.	*output*	[tʌɪpraɪɾə]

Another example of the same argument for imposing ordering on rules, grammars differing only in rule ordering, is examined in **Kiparsky (1982b**: 65–66). German devoices obstruents in coda position (3a) (CHAPTER **69**: FINAL DEVOICING AND FINAL LARYNGEAL NEUTRALIZATION) and simplifies /ŋg/ clusters to [ŋ] (3b). Two of the inflectional forms of the adjective meaning 'long', *lang* and *lange*, contrast in two dialect groups, one showing [laŋk], [laŋə], the other [laŋ], [laŋə], respectively. Application of any of the two rules renders the other rule inapplicable (a case of mutual bleeding, see §2), therefore only the first rule applies in each ordering in every instance:

(3) a. Devoicing $[+\text{obstr}] \rightarrow [-\text{voice}] \;/\; \_\_ \begin{Bmatrix} +\text{C} \\ \# \end{Bmatrix}$

b. *g*-deletion $\text{g} \rightarrow \varnothing \;/\; [+\text{nasal}] \_\_$

c. *Dialect group I*

	/laŋg/	/laŋg+ə/
Devoicing	laŋk	—
g-deletion	—	laŋə

d. *Dialect group II*

	/laŋg/	/laŋg+ə/
g-deletion	laŋ	laŋə
Devoicing	—	—

Rule ordering is closely connected to rule application. As shown 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, i.e. they 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.[3] 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 of ordered rules (4f), (4g):

(4) a. t → ʧ / __ i
b. e → i / __ ʧ

Underlying	*Surface* e. Simultaneous application	f. Rules ordered (a)–(b)	g. Rules ordered (b)–(a)
c. /eti/	eʧi	iʧi	eʧi
d. /teʧ/	tiʧ	tiʧ	ʧiʧ

The problem is now empirical, i.e. the question to ask is whether natural languages have input–output relations like (4c), (4d) to (4e), or rather like (4c), (4d) to (4f) or (4c), (4d) to (4g). In the case of ordered rules, the first rule creates a representation that allows the application of the second rule (feeding). With the ordering (4a) < (4b) ("<" denotes "is ordered before") feeding takes place in /eti/ → eʧi → [iʧi]; with the ordering (4b) < (4a) feeding takes place in /teʧ/ → tiʧ → [ʧiʧ]. Simultaneous application makes these feeding relationships impossible. Thus, since feeding relations are clearly observable in natural languages, the rule-ordering hypothesis receives more support than the hypothesis of simultaneous application.

A similar example can be constructed with two rules, each of which prevents the application of the other if applied first (mutual bleeding; §2). Consider a language with palatalization of velars before /i/ (**CHAPTER 71**: PALATALIZATION) and backing of /i/ to [u] after velars, and the underlying representation /ki/:

(5) a. k → ʧ / __ i
b. e → i / __ ʧ

Underlying	*Surface* d. Simultaneous application	e. Rules ordered (a)–(b)	f. Rules ordered (b)–(a)
c. /ki/	ʧu	ʧi	ku

Under rule ordering, for order (5a) < (5b) we can only apply palatalization (/ki/ → ʧi → (*n/a*)). For order (5b) < (5a), we can only apply backing (/ki/ → ku → (*n/a*)). 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. [ẽ] → [ã]), another raising mid unstressed vowels (e.g. [e] → [i]). They will force /'ẽ/ → ['ã], /e/ → [i], under any application mode. But unstressed /ẽ/ will satisfy 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 /ẽ/ meets the structural description of both rules, two simultaneous contradictory changes must apply to /ẽ/: it has to be lowered *and* raised.

2 Total strict order and other ordering relations

Rule ordering has been a formal property of generative phonology since its beginnings (**Chomsky 1951**; **Chomsky *et al.* 1956**; **Halle 1959**). As defined in *SPE*, the set of rules form a total (or linear) strict order, i.e. the relation "precede" or "is ordered before" relating two rules has the properties (6a)–(6d) (see CHAPTER 34: PRECEDENCE RELATIONS IN PHONOLOGY for formal discussion of a different interpretation of precedence):

(6) Precedence (A < B, or "rule A precedes rule B") is:

a. asymmetric: given any two rules A, B, it is not the case that A precedes B and B precedes A;

b. (hence) irreflexive: given any rule A, it is not the case that A precedes A,

c. transitive: given any three rules A, B, C, if A precedes B and B precedes C, then A precedes C;

d. total (or connected): for any pair of rules A, B, either A precedes B or B precedes A.

Notice that in many cases some rules are not *crucially* ordered: the same surface form will result with the ordering A < B and B < A. Still, we assume that in the generation of the derivation that begins with the underlying form and ends with the surface form they apply in a given order. In *SPE* there were small departures from total strict order, most notably in the case of disjunctive ordering and simultaneous application in the case of infinite rule schemata, both discussed below.

Rule ordering generated a lot of discussion in the 1960s and the beginning of the 1970s. Some of the questions that were asked were whether ordering of processes was justified, whether it could be partially or totally derived from general principles, and what were the possible types of ordering. I examine now some of the proposals that were advanced in this period.

2.1 Eliminating extrinsic ordering

Since in many cases rules are not crucially ordered, a first question that was asked is whether extrinsic ordering could be dispensed with. The ordering between two rules A and B is *extrinsic* if it is imposed language-specifically, as in the Canadian Raising examples. On the other hand, *intrinsic* ordering refers to an ordering imposed by universal properties of grammars and by formal properties of rules. Consider the interaction of /ɣ/-deletion (7a) and Diphthongization (7b) in Finnish, analyzed by **Kiparsky (1968**: 177) as the interaction of two extrinsically ordered rules. But **Koutsoudas *et al.* (1974)** argue that if we assume that rules apply sequentially, but are unordered extrinsically in the sense that they are "simply applied to every representation that satisfies its structural description," we get the right results, no matter which rule we try to apply first ((7a) or (7b)). Consider the derivation of /vee/ and /teɣe/ in (7). /ɣ/-deletion will not be able to apply in the derivation of /vee/ either before or after Diphthongization, and we will get /vee/ → [vie]. In the second example, Diphthongization cannot apply to /teɣe/, but /ɣ/-deletion can, with the result /teɣe/ → tee. Now Diphthongization can apply, deriving the final output, tee → [tie]. Here the ordering /ɣ/-deletion < Diphthongization is not established extrinsically, but it is determined intrinsically by the form of the rules.

(7)

			/vee/	/teɣe/
a.	ɣ	→ Ø / V __ V	—	tee
	ee	→ ie	vie	tie
b.	ee	→ ie	vie	—
	ɣ	→ Ø / V __ V	—	tee
	ee	→ ie	—	tie

Using this and other examples, **Koutsoudas *et al.* (1974)** conclude that rules are unrestricted in their application, every rule applying to every representation that satisfies its structural description, and that no extrinsic ordering needs to be imposed on them. They base their argumentation on four possible situations that derive from the possible orderings in terms of feeding, counterfeeding and counterbleeding relations. I will analyze such relations in more detail in §3; here I only give the basic definitions:

(8) Given rules A, B, where A < B:

a. A feeds B (or A and B are in feeding order) iff A can increase the number of representations to which B can apply;

b. A bleeds B (or A and B are in bleeding order) iff A can decrease the number of representations to which B can apply;

c. B counterfeeds A (or A and B are in counterfeeding order) iff B would feed A if the order were B < A;

d. B counterbleeds A (or A and B are in counterbleeding order) iff B would bleed A if the order were B < A.

Cases of pure feeding order (when it does not involve counterfeeding or counterbleeding at the same time), can be dealt with under these assumptions, as we have seen in (7). The second possible situation is a pure counterbleeding relation. Consider Koutsoudas *et al.*'s example, taken from **Kiparsky (1968**: 199). In certain Low German dialects, post-vocalic voiced stops spirantize. A word like *Tag* 'day' spirantizes the underlying /g/ in the plural [ta:ɣə], but in the singular it spirantizes and devoices by Final Devoicing, yielding [ta:x]. Kiparsky proposes a pair of ordered rules (9a), (9b) that determine the derivation in (9c). But Koutsoudas *et al.* argue that if the rules apply simultaneously the same result is achieved, as shown in (9d): since /g/ is a post-vocalic voiced stop and also a word-final obstruent, it meets the structural description of both rules.

(9) a. Spirantization $\begin{bmatrix}\text{+stop}\\ \text{+voice}\end{bmatrix} \rightarrow [\text{–stop}]$ / V __

b. Devoicing [+obstruent] → [–voice] / __ #

c.		/ta:g/	d.		/ta:g/
	Spirantization	ta:ɣ		Spirantization, Devoicing	ta:x
	Devoicing	ta:x		(applied simultaneously)	

A third set of cases involves two rules that are in both bleeding and counter-bleeding order. One such case is the Latin American Spanish [l ~ j] alternation in pairs like [a'kel] ~ [a'kejos] 'that (MASC SG ~ MASC PL)', analyzed by **Saporta (1965)** with two extrinsically ordered rules. Given underlying forms /akeʎ/, /akeʎos/ the first rule depalatalizes underlying /ʎ/ word-finally; the second turns any remaining /ʎ/ into [j].

(10)		/akeʎ/	/akeʎos/
Final Depalatalization	ʎ → l / __#	akel	—
Delateralization	ʎ → j	—	akejos

Cases like these are dealt with by **Koutsoudas *et al.* (1974)** by invoking a universal principle, very similar to Kiparsky's Elsewhere Condition (see §2.3), the Proper Inclusion Precedence Principle, which determines intrinsic ordering for a certain class of rules:

> (11) *Proper Inclusion Precedence*
> For any representation R, which meets the structural descriptions of each of two rules A and B, A takes applicational precedence over B with respect to R if and only if the structural description of A properly includes the structural description of B.

Given Proper Inclusion Precedence, the rules in (10) can be left (extrinsically) unordered; since the structural description of Final Depalatalization (ʎ# = "any word-final ʎ") is included in the structural description of Delateralization (ʎ = "any ʎ'), Final Depalatalization takes precedence. Notice that this is similar to disjunctive ordering in *SPE*, where a rule containing abbreviatory parentheses like A → B / ___ C(D) stands for two rules, A → B / ___ CD and A → B / ___ C, the first rule applying first and application of one rule excluding application of the other rule. Disjunctive ordering, though, is limited to rules that are adjacent in the ordering.

The fourth set of cases is formed by pairs of rules which are in some other ordering relation. For such cases Koutsoudas *et al.* argue that there are alternative analyses in which no extrinsic ordering is necessary.

The proposal that extrinsic ordering should be dispensed with was also made by Natural Generative Phonology, a theory that began with work by Theo Vennemann and was presented in detail in **Hooper (1976)**. In Natural Generative Phonology, many ordering relations just disappeared, because typical *SPE* phonological rules became either morphophonemic rules or "via rules," rules relating surface representations and hence not taking any part in derivations. In addition, the "No-Ordering Condition" explicitly prohibited extrinsic ordering.

2.2 Multiple application, local ordering, and global rules

Although based on total strict ordering (6), *SPE* allowed for multiple application of a rule in the case of inifinite schemata (**Chomsky and Halle 1968**: 343–348), an abbreviatory convention by which X_n stands for an infinite set of sequences of n or more instances of X. Thus given a rule like C → Ø / ___ C_1# (where C_1 abbreviates the infinite set {C, CC, CCC, …}) and a string … VCCC, both the first and the second C meet the structural description of the rule, which applies simultaneously to both and derives … VC#.

Another case of multiple application of a rule that was introduced after *SPE* is iterative application. In iterative rule application, which is usually directional, after the first structural change takes place, the string is scanned again and if the structural description is met, the rule reapplies, the string is scanned again, etc., until the rule is no longer applicable. Processes like harmony (**chapter 91**: vowel harmony: opaque and transparent vowels; **chapter 118**: turkish vowel harmony; **chapter 123**: hungarian vowel harmony; **chapter 110**: metaphony in romance), stress (**chapter 40**: the foot; **chapter 41**: the representation of word stress; **chapter 39**: stress: phonotactic and phonetic evidence), tone spreading (**chapter 45**: the representation of tone; **chapter 114**: bantu tone), and metrical structure construction are cases of iterative application. Notice that iterative rules do not follow strict ordering because a rule can precede itself, hence the property of irreflexivity in (6b) is abandoned. **Kenstowicz and Kisseberth (1977**: 180–183) present the following case of iterative application. In Gidabal, a long vowel shortens after another long vowel. In a word like /djalum-ba:-da:ŋ-be:/ 'is certainly right on the fish', simultaneous application would derive the wrong form, *[djalum-ba:-daŋ-be], with two vowels shortened, but iterative left-to-right application yields the correct derivation. We see /djalum-ba:-da:ŋ-be:/ → djalum-ba:-daŋ-be: in the first iteration, and no change, because the structural description is not met, in the second iteration. Different versions of iterative application are discussed in **Anderson (1974**: 124–133) and **Kenstowicz and Kisseberth (1977**: 177–195).

Another case in which precedence can be reflexive, against strict order (6), is the case of "anywhere" or "persistent" rules.[4] These are rules that apply in the derivation whenever their structural description is met. A typical case is (re)syllabification. If a rule A establishes some syllabic properties to which rule B is sensitive, and rule B introduces changes, like insertion, deletion, or changes in sonority properties, to which the syllabification rule A is sensitive, A has to apply before and after B. **B. Myers (1991)** argues for several other cases of iterative ordering.

The asymmetric property of strict order (6a) is also abandoned in the case of *local ordering*, a specific type of ordering proposed in **Anderson (1969)** and also discussed in **Anderson (1974)**. He proposes that two rules might have to apply in different orderings to different representations. The ordering in which they apply is the *unmarked order*. Relative to a given representation, feeding is unmarked with respect to a neutral ordering, and bleeding is marked with respect to a neutral ordering, hence the unmarked orderings are feeding and counterbleeding. This proposal stems from the observation in **Kiparsky (1968)** that when a linguistic change consists of a reversal in the ordering of two rules, feeding order tends to be maximized and bleeding order tend to be minimized. To see how this works, consider Icelandic Umlaut and Syncope (**Anderson 1969**, 20–27, **1974**: 141–147) applying to [kœtlum] 'kettle-DAT PL' and [bœggli] 'parcel-DAT SG':

(12) *Local ordering: (a) applies before (b) in (c); (b) applies before (a) in (d)*

a. Umlaut $a \rightarrow œ$ / __C_0 u

b. Syncope $\begin{bmatrix} +\text{syll} \\ -\text{stress} \end{bmatrix} \rightarrow \varnothing$ / C __ C+V

	c.		d.
	/katil+um/		/bagg+ul+e/
Syncope	Ø	Umlaut	œ
Umlaut	œ	Syncope	Ø
output	kœtlum		bœggli

Notice that in (12c) for the representation /katil+um/, Syncope feeds Umlaut, and if the order were the opposite, Umlaut would either feed or bleed Syncope, i.e. order would be neutral. Hence the unmarked order Syncope < Umlaut is chosen. In (12d), for the representation /bagg+ul+e/, the first rule, Umlaut, doesn't either feed or bleed the second rule, Syncope, i.e. the order is neutral. But if the order were the opposite, Syncope would bleed Umlaut. The neutral ordering Umlaut < Syncope is therefore preferred.

There is yet another type of rule that relates to rule application and to ordering, the global rule. A global rule must be applied at the level at which it appears in the ordering, but crucial information for its structural description is ordered elsewhere in the derivation. **Kisseberth and Abasheikh (1975)** present the following case. In Mwini the perfect is formed by adding the suffix /-iːɬ ~ -eːɬ/ to the verb base, as shown in (13). After a base-final sibilant or /ŋ/, the /ɬ/ in the suffix spirantizes to [z] (13b). The final consonant of the verb base undergoes Mutation, which turns coronal and labial stops into [s], /k/ into [ʃ], and a nasal voiced stop sequence into [nz] (13c) (the data are from **Hyman's 1993** analysis of the processes).

(13)	*Root*	*Perfect*	
a.	/jiːb/	jib-iːɬ-e	'he answered'
	/soːm/	som-eːl-e	'he read'
b.	/aːnz/	anz-iːz-e	'he began'
	/t̪oʃ/	t̪oʃ-eːz-e	'he thought'
c.	/fuːŋg/	-fuːnz-il-e	'he closed'
	/pik/	-piʃ-il-e	'he swore'

As shown by the examples in (13c), the vowel in the suffix undergoes a further change: it shortens. But shortening cannot be determined by the base-final sibilant because the derived sibilants in (13b) do not cause shortening. Hence the correct generalization seems to be that shortening applies just in case Mutation has applied, and therefore we need to make reference to another stage in the derivation instead of being restricted to the local stage of application.[5]

The *SPE* formalism 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*, disjunctive ordering derives from abbreviatory devices: … A(B) … corresponds to two disjunctively ordered rules, … AB … and … B …, the first immediately preceding the second, and the application of one rule excluding the application of the other rule. Disjunctive ordering was later derived as an effect of a more general principle, the Elsewhere Condition, originally formulated by **Kiparsky (1973a**: 94). (14) gives the revised formulation in **Kiparsky (1982a**: 136):

(14) Rules A, B apply disjunctively to a form Φ if and only if:

a. The structural description of A (the special rule) properly includes the structural description of B (the general rule).

b. 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 non-high vowels 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 (15c), (15d); lengthened vowel underlined). Otherwise, they are shortened (15e) (see also **CHAPTER 88**: DERIVED ENVIRONMENT EFFECTS):

(15)

a. $\begin{bmatrix} \text{V} \\ -\text{high} \end{bmatrix} \rightarrow \text{V:} \;/\; [\,\_\_\; \text{C}_0\; \text{i}]_{\text{StressFt}}\; \text{V}$

b. $\begin{bmatrix} \text{V} \\ -\text{high} \end{bmatrix} \rightarrow \text{V:} \;/\; [\,\_\_\; \text{C}_0\; \text{V}]_{\text{StressFt}}$

c. *re[medi]al* *[radi]al*
co[loni]al *De[voni]an*

d. *[trivi]al* *di[vini]ty*
[natu]ral *ex[plana]tory*

(15a), (15b) are two rules applying in the general context "disyllabic stress foot with stressed non-high vowel." (15a) 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, –high] C_0 i]$_{\text{StessrFt}}$ V is a subset of [[V, –high] C_0 V]$_{\text{StressFt}}$, (15a) and (15b) are in an Elsewhere relationship: (15a) is the specific rule applying in [[V, –high] C_0 i]$_{\text{StessrFt}}$ V and (15b), the general rule, applies elsewhere in the general common context.

Kiparsky (1982a) makes ingenious use of the Elsewhere Condition to predict blocking of general morphological processes by more specific ones, and strict cyclic effects like non-derived environment blocking (NDEB). An example of the first case is the plural rule in (16c), which applies to the singular form *person*, but shouldn't apply to the lexical plural *people*. Kiparsky assumes that every lexical item L is an identity rule L ↔ L. Since the structural description [*people*]$_{\textit{Noun}, +\text{Pl}}$ in (16a) is a proper subset of the structural description]$_{\text{Noun} +\text{Pl}}$ in (16c), the identity rule applies (vacuously) and blocks the application of the plural rule. The structural description [*person*] $_{\text{Noun} +\text{Pl}}$, on the other hand, is not a lexical item, but the pluralization of the lexical item, so that the plural rule applies regularly to derive *persons*.

(16)

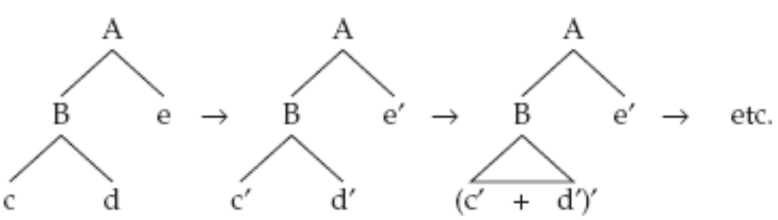
	Rules	*Derivations*	
		[*people*]$_{\text{Noun}, +\text{Pl}}$	[*person*]$_{\text{Noun}, +\text{Pl}}$
a.	[*people*]$_{\text{Noun}, +\text{Pl}}$ ↔ [*people*]$_{\text{Noun}, +\text{Pl}}$	[*people*]$_{\text{Noun}, +\text{Pl}}$	
b.	[*person*] ↔ [*person*]		—
c.	Ø → z / __]$_{\text{Noun}, +\text{Pl}}$	—	[*person*]$_{\text{Noun}, +\text{Pl}}$ /-z/

The Elsewhere Condition can also prevent rule application in non-derived environments (non-derived environment blocking, or NDEB; **CHAPTER 88**: DERIVED ENVIRONMENT EFFECTS), an effect which is also predicted by the Strict Cycle Condition (see next section). Thus the shortening rule (cf. (15b)) shortens [[*san*]*ity*] to [[*săn*]*ity*], because it is derived, but fails to apply to [*nightingale*] (*[*nĭghtingale*]), because it is underived. Again the identity rule [*san*] ↔ [*san*] cannot block shortening, because the structural description of the identity rule, [*san*], and the structural description of shortening, [V C_0 V]$_{\text{StrF}}$, are not in a proper inclusion relation, but in intersection. But [*nightingale*] properly includes [V C_0 V]$_{\text{StressFt}}$. Cyclic effects in later cycles/levels derive from the assumption that the output of every cycle/level is a lexical item, hence an identity rule.[6]

3 Cyclic ordering

In cyclic application the input to the set of rules is a phonological representation organized by constituent structure (**CHAPTER 85**: CYCLICITY). Consider a structure [[c d]$_B$ e]$_A$, which is represented in (17) by a tree diagram; A, B are categories, and d, e, c are phonological strings, matrices, or structures.

(17)



(17) shows a cyclic derivation of $[[c\ d]_B\ e]_A$. The set of 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 is (c' + d')'. 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 *et al.* 1956**) to deal with stress in compounds like *blackboard eraser*, showing primary–tertiary–secondary stress distinctions (stress levels are indicated above the vowel in (18), and as subscripts in the text below):

(18) $\quad \overset{1}{} \quad \overset{3}{} \quad \overset{2}{}$

$[[[\textit{black}]\ [\textit{board}]]_B\ [\textit{eraser}]]_A$

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 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 [*black board*], 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*].

Cyclicity 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 this version of the SCC also in phonology. In (17), for instance, in the second cycle, cycle B, a rule cannot apply to the domain of B if it affects just c'. An actual example is the interaction of Glide Formation and Destressing in Catalan (**Mascaró 1976**: 24–36). Glide Formation applies to post-vocalic unstressed high vowels. In *produirà* 'it will produce' it cannot apply to [[prudu 'i]$_1$'ra]$_2$ at cycle 1, because post-vocalic /i/ is stressed. At cycle 2, a following stress causes destressing of /i/. Therefore at cycle 3, and at later cycles, the sequence /ui/ meets the structural description of the rule; but /ui/ is entirely within cycle 1 and the SCC blocks application, resulting in [pruðui'ra] *[pruðuj'ra].

The SCC was further refined in **Mascaró (1976**: 1–40), as in (19). Case (19b.i) corresponds roughly to the SCC as formulated in **Chomsky (1973)** and used by **Kean (1974)**.

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

a. C applies to the domain $[\ldots]_j$ after having applied to the domain $[\ldots]_{j-1}$, 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 by virtue of) cycle j. This situation obtains if (i), (ii), or (iii) is met:

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

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

iii. 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 (19b.i), derived environments by compounding or syntax (19b.ii), and rule-derived environments (19b.iii). Effects of derived environments on application of processes, irrespective of the theoretical mechanism they derive from, are usually referred to as *derived environment effects* (DEE). We just saw a case, the interaction of glide formation and destressing in Catalan, which falls under (19b.i). The rule of /t/ → [s] assibilation in Finnish illustrates both (19b.i) and (19b.ii). Assibilation (20a) applies in morphologically derived environments like (20c): the structural description [ti] is met by material in the root cycle and in the inflected word cycle. It also applies in rule-derived environments (20d): here the structural description [ti] is met because at its cycle of application, the rule of raising (20b) has created it. But it fails to apply in the non-derived environments (20e), because none of the conditions for proper application in (19b) is met:

(20) a. t → s / __ i
b. e → i / __##
c. /halut-i/ → halus-i 'wanted' cf. halut-a 'to want'
d. /vete/ → veti → vesi 'water (NOM SG)' cf. vete-næ 'water (ESS SG)'
e. tila 'place, room'
æiti 'mother'
itikka 'mosquito'

An instance of (19b.ii) is the application of glide formation in Central Catalan to vowels of different words. As we have just seen, glide formation applying to post-vocalic high vowels is blocked in *produirà* [[pruð'i]$_1$ $^{\text{i}}$ra]$_2$. Consider now *produirà oxidació* 'it will produce oxidation':

(21)

	[[[pɾuðu'i]$_1$ 'ɾa]$_2$	[[['ɔksið]$_1$ 'a]$_2$ 'sjo]$_3$]$_4$
Cycle 2	[[pɾuðu'i]$_1$'ɾa]$_2$	[['ɔksið]$_1$ 'a]$_2$
	[[pɾuðui]$_1$ 'ɾa]$_2$	[[uksið]$_1$ 'a]$_2$
Cycle 3		[[[uksið]$_1$ ə]$_2$ 'sjo]$_3$
		[[[uksið]$_1$ ə]$_2$ 'sjo]$_3$
Cycle 4	[[[pɾuðui]$_1$ 'ɾa]$_2$	[[[uksið]$_1$ ə]$_2$ 'sjo]$_3$]$_4$
	[[[pɾuðui]$_1$ 'ɾa]$_2$	[[[wksið]$_1$ ə]$_2$ 'sjo]$_3$]$_4$

In the second word, at cycle 2, the initial /$^{\text{i}}$ɔ/ is destressed by a following stress and becomes [u] by a rule of vowel reduction. At cycle 4, the sequence /au/ meets the structural description of glide formation and the SCC does not block Glide Formation, the application being proper by (19b.ii), because /au/ is not within the domain of a single previous cycle: /a/ is in cycle 2; /u/ is in cycle 3. Hence the rule applies, yielding ['aw].

It was assumed that the SCC applied to cyclic, obligatory neutralization rules, and dealt with DEEs. These were previously accounted for by the Alternation Condition proposed by **Kiparsky (1973b**: 65), according to which "neutralization processes apply only to derived forms … [i.e.] 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 post-lexical phonology, which correspond to cycles and to the effect of the Elsewhere Condition from which DEEs are derived. In Stratal Optimality Theory (see **CHAPTER 85**: CYCLICITY), cycles correspond to strata to which Gen and Eval apply successively. Within OT, output-output faithfulness constraints (**Benua 1997**) ensure similarity of larger constituents to its inner components. Strict cycle effects (DEEs) are also obtained by local conjunction of markedness and faithfulness constraints (**Łubowicz 2002**). To see how DEEs are derived from local conjunction, consider the interaction of Velar Palatalization and Spirantization in Polish (**Łubowicz 2002**: §3). We find the following descriptive generalizations. Spirantization applies to rule-derived [ʥ] (22a), but not to underlying /ʥ/ (22b); similarly, in (22c) Velar Palatalization applies only to morphologically derived velar + [e i] sequences.

(22)

	a.	b.	c.
	/rog-ek/ 'horn'	/banʤ-o/ 'banjo'	/xemik-ek/ 'chemist-DIM'
Velar Palatalization	roʤ-ek	—	xemiʧ-ek *blocked in* xe
Spirantization of /ʤ/	roʒ-ek	*blocked*	—
output	roʒ-ek	banʤ-o *banʒ-o	xemiʧ-ek *ʃemiʧ-ek

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

For morphologically derived environments, as in the example in (22c), Łubowicz uses conjunction of markedness and ANCHOR. Velar palatalization applies to the morphologically derived sequence [k-e], but not to the non-derived sequence /xe/ in /xemik-ek/. Since the velar /k/ is stem-final, but not syllable-final, in [xe.mi.k]$_{\text{Stem}}$-ek, the sequence k]$_{\text{Stem}}$-e will violate R-ANCHOR(Stem, s), 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, s)]$_{\text{D}}$. But since morphologically underived /xe/ satisfies R-ANCHOR(Stem, s), the conjunction will be satisfied in this case and palatalization will not take place.

4 Rule interaction, ordering, and applicability: Feeding and bleeding

In a system in which rules are ordered, rules can interact: both the applicability and the result of application of a rule can depend on the application of previous rules. The notions of *feeding* and *bleeding* that I made reference to in §2 were introduced by **Kiparsky (1968)** in order to explain the direction of linguistic change. These concepts have been widely used since. In this section I examine them in some detail.

Since it is not uncommon to detect terminological inadequacies in the literature, in order to 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 presupposed. 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). The terms *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 a *feeding* or *bleeding relation* can be determined by mere inspection of the rules.[7] 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 make use of the predicates *p-feed* and *p-bleed* (“p” for “potentially”) in the case of feeding/bleeding *relations*. (23) provides an illustration using our previous German example (3):

(23) *German, group II* (*g*-deletion < Devoicing)

a.

Feeding/bleeding relation	A p-feeds/ p-bleeds B	devoicing p-bleeds *g*-deletion *g*-deletion p-bleeds devoicing
Feeding/bleeding order	A feeds/ bleeds B	*g*-deletion bleeds devoicing

b.

Devoicing	$[\text{obstr}] \rightarrow [-\text{voice}] \ / \ \_\_ \begin{Bmatrix} +\text{C} \\ \# \end{Bmatrix}$
g-deletion	$\text{g} \rightarrow \varnothing \ / \ [+\text{nasal}] \ \_\_$

c. *Dialect group II*

	/lang/
g-deletion	laŋ
devoicing	—

Devoicing p-bleeds *g*-deletion by devoicing *g* in [+nasal] + *g*, and *g*-deletion p-bleeds devoicing by deleting *g* in the same context. Given the ordering *g*-deletion < devoicing, *g*-deletion bleeds devoicing, as shown in the derivation in (23).

Feeding and bleeding relations can be formally defined as follows:

(24) *Feeding and bleeding relations*

a. Rule A is in feeding relation with respect to 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.

b. Rule A is in bleeding relation with respect to 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 (24) “apply” is usually interpreted as “apply non-vacuously.” In the German example in (3), in dialect group I, Devoicing bleeds *g*-deletion, (/laŋg/ → /laŋk/ → (*n/a*)). But for the word *Bank* ‘bank’, whose derivation is /baŋk/ → (vacuous devoicing) baŋk → (*n/a*), we don't want to say that Devoicing bleeds *g*-deletion, because the input to Devoicing didn't meet its structural description. **Kiparsky's (1968)** terms "creates" and "removes," cited 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 which presuppose both feeding/bleeding relations and the specific ordering A < B (i.e. A precedes B) in the grammar. Most definitions are formulated for cases in which A immediately 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 with respect to B iff A < B and A p-feeds/bleeds B. For the general case the definitions have to be refined as follows:

(25) *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 with respect to 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 with respect to 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 immediately precedes B or in cases where intermediate rules don't interact we get derivations like those in (26): (26a.i) is in feeding order with respect to (26a.ii) because the second rule (26a.ii) wouldn't apply to AQ, but applies to BQ, the output of the first rule (26a.i); (26b.i) is in bleeding order with respect to (26b.ii) because the second rule (26b.ii) would apply to AQ, but doesn't apply to BQ, the output of the first rule (26b.ii).

(26) a. *Feeding order*
(No intervening interacting rules)

		AQ
i.	A → B / __ Q	BQ
ii.	Q → R / B __	BR

b. *Bleeding order*

		AQ
i.	A → B / __ Q	BQ
ii.	Q → R / A __	—

The case of feeding order for two adjacent rules can be illustrated with the interaction of /æɬ/ → [a:] and Umlaut in a group of Swiss German dialects (**Kiparsky 1982b**: 190). Bleeding order, also for adjacent rules, can be illustrated with our earlier example (2e)–(2h), Canadian Raising, in the word *writer* in dialect B:

(27) a. *Feeding: Swiss German (dialect group I)*

		/æi̯-li/	'egg-DIM'
i.	æi̯ → aː / __ {C, ##}	aːli	
ii.	Umlaut (fronting)	æli	(does not apply non-vacuously to /æi̯-li/)

b. *Bleeding: Canadian Raising (dialect B)*

		/raɪtə/	
i.	t → ɾ / V __ V	raɪɾə	
ii.	a → ʌ / __ [C, −voice]	—	(applies non-vacuously to /raɪtə/)

Consider now the cases with interacting rules intervening between (i) and (ii) that motivate the definitions in (25). (28a) exemplifies feeding cases and (28b) bleeding cases. The rules (i) and (ii) are the rules in feeding/bleeding relation; (iii) is the intervening rule.

(28) a. *Feeding order* (Intervening interacting rules)

		QA
i.	Q → R / __ A	RA
iii.	A → B / R __	RB
ii.	A → C / R __	—

b. *Bleeding order*

		QA
i.	Q → R / __ A	RA
iii.	R → Q / __ A	QA
ii.	A → B / Q __	QB

In the feeding example, rule (i) p-feeds rule (ii) and precedes (ii), but given conditions (25a.ii) and (25b.ii), it does not feed rule (ii), because some rule ordered between them, namely (iii), undoes the change that caused the feeding (it bleeds rule (ii)). In terms of the definitions in (25), there are representations between the two rules, in particular the input to rule (ii), to which the second rule cannot apply. Similarly, in the bleeding example, rule (i) p-bleeds rule ii. and would indeed bleed rule (ii), if it were not for (iii), which feeds rule (ii).

Of course a pair of rules can show non-feeding or non-bleeding interactions like those in (28) in some derivations, but feeding or bleeding interactions in other derivations. As already indicated in note 7, 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* ['kan, tekst] or *Ahab* ['Eʊ,hæb]. But after a light syllable the second stress is removed (the "Arab rule"; **Ross 1972**), as in *Arab* ['ærəb], and the destressed vowel reduces to [ə]. Stress bleeds vowel reduction, but in the derivation of *Arab* destressing undoes the bleeding (/ærəb/ → 'æ,ræb → 'æræb → ['ærəb]). Here we must say that stress bleeds reduction, because there are derivations that show actual bleeding, as in ['Eʊ'hæb], 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.[8]

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 (25), 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 (**Pullum 1976**; **McCarthy 2003**), which are derivations in which a rule reverses the action of a previous rule, e.g. … A … → … B … → … A … Consider the derivations in (29), which contain a Duke of York subderivation, highlighted in bold (notice that (28b) above is also an instance of a Duke of York derivation):

(29) a. *Feeding order* b. *Bleeding order*
(Intervening interacting rules; Duke of York derivations)

		QA			QA
i.	Q → R / __ A	**RA**	i.	Q → R / __ A	**RA**
iii.	A → B / R __	**RB**	iii.	R → Q / __ A	**QA**
iv.	B → A / R __	**RA**	iv.	A → B / Q __	QB
ii.	A → C / R __	BC	ii.	A → C / Q __	__

In (29a) (i) p-feeds and precedes (ii), and (ii) does apply, but (i) does not feed (ii), because there is an intermediate representation to which the second rule would not apply, namely RB, created by (iii). In fact it is the other intervening rule, (iv), that now feeds (i). Similarly, in (29b) (i) p-bleeds the last rule (ii), but it does not bleed it, even if the rule does not apply, because of the intermediate representation QA created by (iii), to which the rule would apply. Here it is rule (iv) which actually bleeds (ii).

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/bleeding relation, and yet this relation is intuitively correct. Consider a case like (30), in which glide formation, vowel reduction, and destressing interact in Central Catalan. (30a)–(30c) show that glide formation affects post-vocalic high unstressed vowels (30a), but not non-high vowels (30b), where *Osiris* is a lexical exception to vowel reduction, or stressed vowels (30c). It also affects high unstressed vowels that are not underlying, as in (30d).

(30)

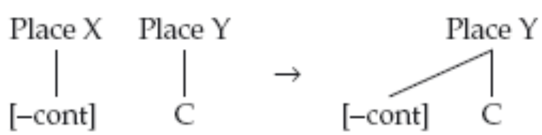
	a. ser['a] [u]mid 'it will be wet'	b. ser['a] [o]siris 'it will be Osiris'
Destressing		
Vowel Reduction		
Glide Formation	'a w	
output	ser['a] [w]mid	ser['a] [o]siris

	c. ser['a] ['u]til 'it will be useful'	d. ser['a] ['ɔ]ciós 'it will be idle'
Destressing		'a ɔ
Vowel Reduction		'a u
Glide Formation		'a w
output	ser['a] ['u]til	ser['a] [w]ciós

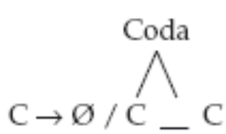
Notice now that the structural description V V[+high, -stress] is met in (30d), because vowel reduction has turned [D] into a high vowel, but also because destressing has created the other condition for gliding. In such a case we want to say that these two rules jointly feed glide formation. The definitions in (25) can be changed accordingly, to meet such situations.

Notice also that a rule can stand in both feeding and bleeding order with respect to another rule. In Majorcan 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 (31), 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.

(31) a. *Place assimilation*



b. *Cluster simplification*



c.

	Bleeding	*Feeding*
input	/'bujd 'trens/ 'I empty trains'	/'trens 'bujds/ 'empty trains'
Cluster simplification	'buj 'trens	'tren 'bujds
Place assimilation	—	'trem 'bujds
output	'buj 'trens	'trem 'bujts

It should be observed that the fact that two rules A, B do not have a feeding or bleeding interaction does not mean that they don't interact. In (32) rule (a) deletes final consonants, while (b) stresses the final syllable if it is heavy, otherwise the penult. Different orderings give different results, but the interaction isn't either a feeding or a bleeding relation.

(32) a. C → Ø / __ ## /satopek/ $\xrightarrow{a}$ satope $\xrightarrow{b}$ sa'tope
b. V → [+stress] / __ C(V)## /satopek/ $\xrightarrow{b}$ sato'pek $\xrightarrow{a}$ sato'pe

I will now examine *counterfeeding* and *counterbleeding*. These notions refer only to rules that are in a specific order (potential situations don't make sense in this context). Basically, a counterfeeding/bleeding order is an order that would be feeding/bleeding if the order of the rules were reversed. Since there is some confusion in the use of the predicates, I will follow the practice in **Koutsoudas *et al.* (1974)** and use as subject of "counterfeed/counterbleed" the second rule in the ordering, e.g. B counterfeeds A means that A < B, and B would feed A if B < A.

(33) a. A and B *are in counterfeeding order* (B *counterfeeds* A) in grammar G iff
   i. A < B
   ii. B p-feeds A

b. A and B *are in counterbleeding order* (B *counterbleeds* A) in grammar G iff
   i. A < B
   ii. B p-bleeds A

Counterfeeding order can be illustrated with the same processes of Swiss German presented in (27) in another dialect group, group II, which shows the opposite ordering. Counterbleeding is illustrated with Canadian Raising in dialect A (2a):

(34) a. *Counterfeeding: Swiss German (dialect group II)*

		/æi̯-li/ 'egg-DIM'
i.	Umlaut (fronting)	—
ii.	æi̯ → aː / __ $\left\{\begin{array}{c} C \\ \#\# \end{array}\right\}$	aːli

The opposite, feeding ordering would yield /æi̯-li/ → aː-li → æː-li

b. *Counterbleeding: Canadian Raising (dialect A)*

		/raɪtə/
i.	a → ʌ / __ [C, −voice]	rʌɪtə
ii.	t → ɾ / V __ V	rʌɪɾə

The opposite, bleeding ordering would yield /raɪtə/ → raɪɾə → (*n/a*)

"Counter" orderings have important properties. Assume the simple case where rules A and B are adjacent, and A < B. Since in feeding order (B < A) there must be at least one input I such that B is applicable to I, A is not applicable to I, and A is applicable to the output of B (35a), it follows that in the corresponding counterfeeding order where A < B there must be an input (namely I) to which the first rule, now A, does not apply and to which the second rule, now B, applies (35b). Hence the generalization expressed by A 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 output of B, usually the surface representation. In the bleeding order B < A there must be by definition at least one input I such that both A and B are applicable to I (giving different results, I′ and I′, respectively), and A is not applicable to the output of B (35c). It follows that in the corresponding counter-bleeding order A < B there can be an input to which the first rule, now A, applies and to which the second rule, now B, might apply and change the context of application of the first rule (35d). 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 of A, usually the surface representation.

(35)

a. *Feeding*		b. *Counterfeeding*		c. *Bleeding*		d. *Counterbleeding*	
	I		I		I		I
B.	I′	A.	—	B.	I′	A.	I″
A.	I″	B.	I‴	A.	—	B.	— / I‴

In our previous example in (34) for counterfeeding in Swiss German, the fronting dictated by Umlaut is not true in the surface form [aːli]; for the counterbleeding in Canadian Raising, dialect A, the fact that [ʌ] derived from /a/ appears before voiceless consonants is not apparent in the surface form [rʌɪɾə].

It is also important to notice that the existential quantification in the definitions in (25) of feeding and bleeding orders (hence also of counterfeeding and counter-bleeding 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 (25a.ii) that there be an input I whose derivation D meets the conditions required in (25a.ii) does not prevent the existence of another input I' that meets the condition (25b.ii) for bleeding. Hence A can both feed and bleed B (and B can both counterfeed and counterbleed A).

5 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 *opaque* rule interactions. If we compare a standard serial theory like *SPE* with a parallel theory based on constraints like Optimality Theory (OT), pure feeding and pure bleeding order effects (i.e. those that are not also counterfeeding or counterbleeding) are *transparent* interactions and can be derived from both. Consider the well-known case of *e*-raising and /t/ → [s] interaction in Finnish (**Kiparsky 1973b**: 166–172), partially repeated from (20):

(36) vete 'water-NOM SG' halut-i 'wanted'

a.	e → i / __##	veti	—
b.	t → s / __ i	vesi	halus-i

Because both (36a) and (36b) are statements that are true of surface forms, constraints of the form *e##, *ti, dominating conflicting faithfulness constraints, together with other constraints determining the choice of [i] and [s], will derive the output of /vesi/, /halus-i/.

But counterfeeding and counterbleeding are opaque interactions and cause problems for a parallel approach. A process (37a) is opaque (**Kiparsky 1973b**: 79) to the extent that there are phonetic forms in (37b) or (37c); otherwise it is transparent. The derivations (37d) and (37e) illustrate (37b) and (37c), respectively.

(37) a. *Rule:* A → B / C __ D

Opaque surface forms

b. A in the environment C __ D

c. B derived by (a) in an environment different from C __ D

d.		/EAD/
	A → B / C __ D	—
	E → C / __ A	CAD
e.		/CAD/
	A → B / C __ D	CBD
	C → E / __ A	EBD

In (37d) the generalization "A does not appear in C ___ D; B appears instead" expressed by (37a) is not surface-true; the rule *underapplies* with respect to surface representations. In (37e) the generalization "underlying (or intermediate) A is represented by B in C __ D" is not true of the derivation, it is not surface-apparent; the rule *overapplies* with respect to surface representations, since it applies outside its environment. To illustrate with a real example, consider counterfeeding in Madurese (Austronesian, Indonesia) (**McCarthy 2002**: 174–175). Nasality spreads rightwards onto following vowels, but is blocked by oral consonants, and voiced stops delete after a nasal (CHAPTER **78**: NASAL HARMONY):

(38) a.		/naŋgaʔ/	/naŋaʔ/
b.	V → [+nas] / N __	nãŋgaʔ	nãŋãʔ
c.	/b d g/ → Ø / N __	nãŋaʔ	—
	Surface	[nãŋaʔ]	[nãŋãʔ]

In the first derivation, rule (38c) has deleted an oral consonant and has thus partially changed the context of application of rule (38b); rule (38b) underapplies, because if it did apply to the surface representation it would nasalize the second vowel, *[nãã']. The generalization that a nasal vowel nasalizes following vowels across non-oral consonants is not surface-true.

Such an opaque interaction is derivable in an ordered rule system, but not in a system in which markedness generalizations are about surface forms. Consider now a model like OT. For an input /nãã/ (cf. the second derivation in (38)), the constraint hierarchy must favor candidate [nãã'] over candidate *[nãã'] (nasalization spreads across non-oral consonants). Therefore it will also favor the non-opaque candidate *[nãã'] over candidate [nãã'] if the input is /nãŋga'/. Similar considerations apply to counterbleeding opacity. Consider our earlier example, Canadian Raising in dialect A. The change /Aʊt/ → [ʌɪt] does not appear as such in the phonetic representation of *writer*, because the second rule has modified the result of the change, turning the triggering voiceless /t/ into [ɾ]:

(39) a.		/raɪtə/		/taɪp/
b.	aɪ → ʌɪ / __ [C, –voice]	rʌɪtə		tʌɪp
c.	t → ɾ / V __ V	rʌɪɾə	(ʌɪɾ, *not* ʌɪt)	tʌɪp

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

(40)

/raɪtə/	*aɪ[C, −voice]	*VtV	Faith[aɪ]	Faith[t]
a. raɪtə	*!	*		
b. rʌɪtə		*!	*	
☞ c. raɪɾə				*
d. rʌɪɾə			*!	*

6 Conclusion

Rules are generalizations about the distribution of sound in natural languages. Rule ordering is a specific theory about how these generalizations interact to derive a surface representation. The intensive study of many phonological systems using rule ordering has not only produced a rich body of descriptive work, but has also unveiled many deep properties of phonological systems and many theoretical problems that go beyond the model that generated them. When the problem of the theoretical status of phonology was first addressed seriously, it was immediately realized that phonological generalizations could not have two properties at the same time: they could not be absolute generalizations and generalizations about the surface representation. In other words, they could not map a lexical representation to a surface representation in one step (simultaneous rule application), as illustrated by Joos's paradox discussed at the beginning of §1. The response to this fact was that the requirement that generalizations be true of surface representations should be abandoned, and hence that phonological processes had to be ordered. The conviction of many present-day phonologists that the right response is to abandon the other requirement, i.e. that generalizations be absolute, and keep the idea that they apply to surface representations, has been made possible by many decades of work in a framework based on rule ordering. Even if many things have changed since the days in which a phonological description could be based on a system with a depth of ordering of 20 or 30 (i.e. 20 or 30 rules that had to be linearly ordered),[9] serial approaches haven't achieved a total elimination of ordering through mechanisms like the ones described in §2.3. At the same time, many of the properties of phonological systems that have been discovered as the result of work on rule ordering – the existence of opacity, disjunctivity as predicted by the Elsewhere Condition, derived environment effects, and many morphology-phonology interactions – are still important problems that will stimulate further research, for both serial and parallel approaches.

Notes

1 The opaque case (dialect A) had already been discussed by **Harris (1951**: 70–71) and **Chomsky (1962**: 156–157).

2 I transcribe the first vowel of the diphthong as [ʌ], and the voiced *t* as [ɾ], following **Chambers (1973, 2006)**; Joos's phonetic description is slightly different (basically [ɾ] and [d], respectively). Canadian Raising has generated a great deal of discussion. **Kaye (1990)** casts some doubts on the existence of dialect B, which are not clearly formulated. **Mielke *et al.* (2003)** claim that the difference has been phonemicized, e.g. as /nʌjf/ *vs.* /najv/, but **Idsardi (2006)** argues convincingly that there are actual alternations.

3 **Chomsky and Halle's (1968)** example consists of the rules B → X / ___ Y and A → Y / ___ X and the input representations /ABY/ and /BAX/.

4 The term "persistent rule" is due to **Chafe (1968)**.

5 **Hyman (1993)** discusses the Mwini case and the alternatives to a global analysis.

6 Given Optimality Theory, some properties of Elsewhere-type interactions follow as a theorem, Pãnini's Theorem on Contraint-ranking (PTC) (**Prince and Smolensky 1993**: §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 non-vacuously 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 applies) then S is not active on *i*. For a clear and interesting discussion of the relation between the Elsewhere Condition and PTC, see **Baković (2006)**; also **Prince (1997)**.

7 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 … Vl]$_{Coda}$## from / … Vʎ]$_{Coda}$##/, to which B is applicable. But in a language with a single lateral *l*, the feeding interaction will never take place. In such cases, in order to avoid terminological ambiguities 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/is not actually instantiated in that particular derivation.

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

9 "In the segment of the phonological component for Modern Hebrew presented in **Chomsky (1951)**, a depth of ordering that reaches the range of twenty to thirty is demonstrated and this is surely an underestimate" (**Chomsky 1964**: 71).

REFERENCES

Anderson, Stephen R. 1969. *West Scandinavian vowel systems and the ordering of phonological rules*. Ph.D. dissertation, MIT.

Anderson, Stephen R. 1974. *The organization of phonology*. New York: Academic Press.

Anderson, Stephen R. & Paul Kiparsky (eds.) 1973. *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston.

Baković, Eric. 2006. Elsewhere effects in Optimality Theory. In Eric Baković, Junko Itô & John J. McCarthy (eds.) *Wondering at the natural fecundity of things: Essays in honor of Alan Prince*, 23–70. Santa Cruz: Linguistics Research Center, University of California, Santa Cruz. Available at **http://repositories.cdlib.org/lrc/prince/4**.

Benua, Laura. 1997. *Transderivational identity: Phonological relations between words*. Ph.D. dissertation, University of Massachusetts, Amherst. Published 2000, New York: Garland.

Chafe, Wallace L. 1968. The ordering of phonological rules. *International Journal of American Linguistics* (34) . 115–136.

Chambers, J. K. 1973. Canadian Raising. *Canadian Journal of Linguistics* (18) . 131–135.

Chambers, J. K. (ed.) 1975. *Canadian English: Origins and structures*. Toronto: Methuen.

Chambers, J. K. 2006. Canadian Raising retrospect and prospect. *Canadian Journal of Linguistics* (51) . 105–118.

Chomsky, Noam. 1951. *Morphophonemics of modern Hebrew*. M.A. thesis, University of Pennsylvania. Published 1979, New York: Garland.

Chomsky, Noam. 1962. A transformational approach to syntax. In Archibald A. Hill (ed.) *Proceedings of the 3rd Texas Conference on Problems of Linguistic Analysis in English*, 124–158. Austin: University of Texas Press.

Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague & Paris: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1973. *Conditions on transformations*. In Anderson & Kiparsky (1973). 232–286.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Chomsky, Noam, Morris Halle & Fred Lukoff. 1956. On accent and juncture in English. In Morris Halle, Horace Lunt, Hugh MacLean & Cornelis van Schooneveld (eds.) *For Roman Jakobson: Essays on the occasion of his sixtieth birthday*, 65–80. The Hague: Mouton.

Halle, Morris. 1959. *The sound pattern of Russian: A linguistic and acoustical investigation*. The Hague: Mouton.

Halle, Morris. 1962. Phonology in generative grammar. *Word* (18) . 54–72.

Halle, Morris. 1995. Comments on Luigi Burzio's “The rise of Optimality Theory.” *Glot International* (1) . 9–10.

Harris, Zellig S. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.

Hooper, Joan B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.

Hyman, Larry M. 1993. Problems for rule ordering in Bantu: Two Bantu test cases. In John A. Goldsmith (ed.) *The last phonological rule: Reflections on constraints and derivations*, 195–222. Chicago & London: University of Chicago Press.

Idsardi, William J. 2006. Canadian Raising, opacity, and rephonemicization. *Canadian Journal of Linguistics* (51) . 119–126.

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 Mascaró & Marina Nespor (eds.) *Grammar in progress: Glow essays for Henk van Riemsdijk*, 259–263. Dordrecht: Foris.

Kean, Marie-Louise. 1974. The strict cycle in phonology. *Linguistic Inquiry* (5) . 179–203.

Kenstowicz, Michael & Charles W. Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In Emmon Bach & Robert T. Harms (eds.) *Universals in linguistic theory*, 171–202. New York: Holt, Rinehart & Winston. Reprinted In Kiparsky (1982b), 13–43.

Kiparsky, Paul. 1973a. *“Elsewhere” in phonology*. In Anderson & Kiparsky (1973). 93–106.

Kiparsky, Paul. 1973b. Phonological representations: Abstractness, opacity, and global rules. In Osamu, Fujimura (ed.) *Three dimensions of linguistic theory*, 57–86. Tokyo: Taikusha.

Kiparsky, Paul. 1982a. Lexical morphology and phonology. In Linguistic Society of Korea (ed.) *Linguistics in the morning calm*, 3–91. Seoul: Hanshin.

Kiparsky, Paul. 1982b. *Explanation in phonology*. Dordrecht: Foris.

Kisseberth, Charles W. & Mohammad Imam Abasheikh. 1975. The perfective stem in Chimwi:ni and global rules. *Studies in African Linguistics* (6) . 249–266.

Koutsoudas, Andreas, Gerald Sanders & Craig Noll. 1974. The application of phonological rules. *Language* (50) . 1–28.

Łubowicz, 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. Published 1978, Indiana University Linguistics Club.

McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* (16) . 331–399.

McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. 2003. Sympathy, cumulativity, and the Duke-of-York gambit. In Caroline Féry & Ruben van de Vijver (eds.) *The syllable in Optimality Theory*, 23–76. Cambridge: Cambridge University Press.

Mielke, Jeff, Mike Armstrong & Elizabeth Hume. 2003. Looking through opacity. *Theoretical Linguistics* (29) . 123–139.

Myers, Scott. 1987. Vowel shortening in English. *Natural Language and Linguistic Theory* (5) . 485–518.

Myers, Scott. 1991. Persistent rules. *Linguistic Inquiry* (22) . 315–344.

Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper & Row.

Prince, Alan. 1997. Elsewhere and otherwise. *Glot International* (2) . 23–24 (ROA-217).

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Unpublished ms., Rutgers University & University of Colorado, Boulder. Published 2004, Malden, MA & Oxford: Blackwell.

Pullum, Geoffrey K. 1976. The Duke of York gambit. *Journal of Linguistics* (12) . 83–102.

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

Saporta, Sol. 1965. Ordered rules, dialect differences, and historical processes. *Language* (41) . 218–224.

Whitney, William Dwight. 1889. *Sanskrit grammar, including both the classical language, and the older dialects of Veda and Brahmana*. 2nd edn. Cambridge, MA: Harvard University Press. Reprinted 1975, Cambridge, MA: Harvard University Press.

Cite this article

Mascaró, Joan. "Rule Ordering." *The Blackwell Companion to Phonology*. van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds). Blackwell Publishing, 2011. Blackwell Reference Online. 18 January 2012 <http://www.companiontophonology.com/subscriber/tocnode?id=g9781405184236_chunk_g978140518423676>