5:

DERIVATIONAL SEQUENCES AND HARMONY SCALES

Derivational sequences and harmony scales are collections of relationships between phonological structures that are used to pick out a grammatical surface representation: a surface representation is the final form in a sequence in derivational phonology; it is the optimal form on a scale of relative harmony in optimality phonology. The two approaches would match still further if traversing along the steps of a derivational sequence to the end were consistent with traversing through increments in harmony up to the peak.

We formulate this possible analogy between derivation and harmony in 5.1, and analyse the extent to which it holds in 5.2. In 5.3, we show that derivational steps which contradict harmony are ruled out by adding a strict economy condition on derivation length. However, this is too strong, ruling out other derivations for which there is evidence in Slavic languages, which do not contradict harmony. In an appendix, we examine how to foreclose the possibility of harmony scales with multiple optimal members, since sequences do not have multiple endpoints.

5.1 The Derivation-Harmony Analogy

5.1.1 Paths and Landscapes

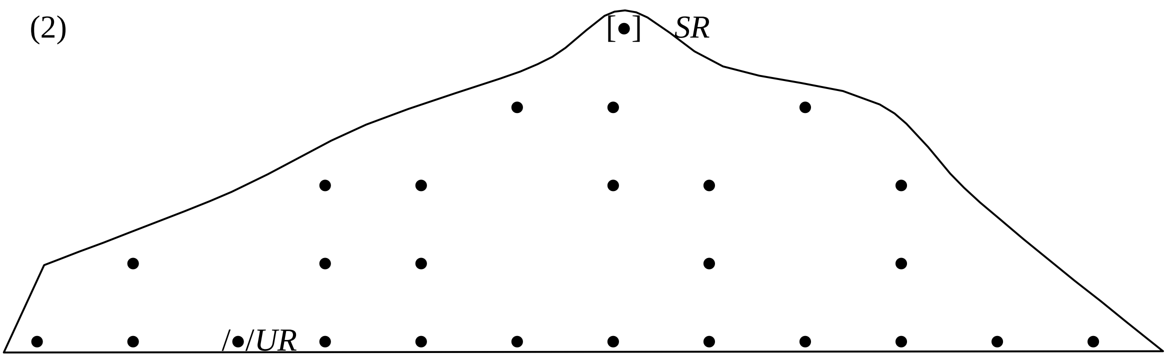
It is noteworthy that in both derivational and optimality phonology, the relation between an underlying form and its surface form is mediated by some wider system of relations between representational structures. In the derivational framework, a sequence of representations is constructed starting from the underlying representation. The surface form is the final form of the sequence constructed. This is illustrated in (1). Representations are shown as a collection of •'s residing in some space, and a derivation is a path through that space, from representation to representation, starting with the underlying representation (*UR*) and leading to the surface representation (*SR*):

(1) /•/ → • → • → • → • → • → [•]

UR *SR*

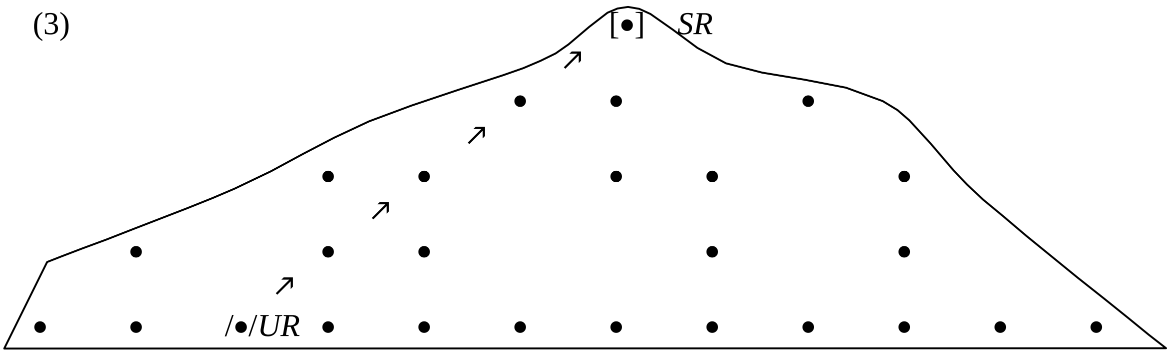
In the optimality framework, each underlying representation is associated with a harmony ordering of all structures under their possible correspondences, and the surface form is the optimal form, the maximally harmonic candidate. This is illustrated in (2). One might think of the space of all possible structures being 'landscaped' by a rating of harmony, with the optimal form residing at the highest peak.

(2)



We can compare these two pictures. The surface representation, for example, is in both cases found in a privileged position: the final member of the derivational path or the peak of the harmony landscape; the fact that there are no further members to the sequence after *SR* correlates with the fact that there is no form more harmonic than *SR*. When we superimpose the derivational path (1) and harmony landscape (2) pictures in (3), the result is a path which tends to *rise* towards the peak:

(3)



This presents us with an analogy between **succession** through the derivational sequence and **incrementation** up the harmony scale, and suggests the conjecture that a derivation P1→P2→... →Pn-1→Pn be matched by the harmonies P1≺P2≺...≺Pn-1≺Pn in a corresponding harmony evaluation.

Note that (3) is a comparison across *two* theories. Some have considered putting together derivation and harmony within a single theory: "a rule *applies* if and only if its effect is to *increase* the well-formedness of the representation" (Goldsmith 1995:7 my italics, cf. Sommerstein 1974, Goldsmith 1990:318ff, 1993). It is true that the comparison we are undertaking bears an anatomical similarity with such theories of "harmonic rule application", but in theory *comparison*, we have derivation and harmony as devices belonging to two separate theories, and we are testing an apparent similarity between those theories. Thus, our construal of a derivational path 'rising' through a landscape resides in a *metatheoretical* frame, articulating a possible correspondence *across* theories. As far as we are concerned here, the fact that other phonologists have thought to place derivation and harmony alongside one another only lends additional support to the venture. In the next two sections, we undergird the analogy formally.

5.1.2 Structures and Candidates

In order for the analogy to make sense in formal terms, we must take care over the harmony relation. Strictly speaking, harmony discriminates between *candidates*, and it was shown in 2.3 that the candidates evaluated are not merely structures, but include a correspondence relation to the input structure. In this they differ from the members of a derivational sequence, which are structures. However, it is possible to resolve this difference.

To think of harmony as a relation between structures is a simplificatory move which is often useful because Markedness constraints - including ONSET, NOCOMPLEX, etc. - are constraints whose evaluation of the candidate focusses entirely on the potential output structure itself. Faithfulness constraints, however, do not focus entirely on the output structure, but evaluate the whole input-output relation. Thus in the tableau (4) below, input/output constraint MAX discriminates between candidates that share output *.ba.* if differing numbers of input elements have correspondents in the output, and similarly between candidates that share *.a.*, while the output constraint ONSET evaluates all instances of *.ba.* identically, and all instances of *.a.* identically, though it does discriminate between different outputs *.ba.* and *.a.* .

(4)

/b_1a_2/	ONSET	MAX
.b_1a_2.		
.b a_2.		*
.a_2.	*	*
.a .	*	**

It is possible to think of markedness constraints as discriminating among structures themselves, *.ba.*, *.a.* etc., as well as the candidates in which they are contained. This leads to the property of **Harmonic Ascent** (Moreton 1999): the optimal output structure is either equal to the input structure or is a more unmarked structure than the input structure when assessed against the subhierarchy of markedness constraints. For example, in (4), *.a.* is not a viable output from *.ba.*, being *more* marked than *.ba.* itself. One way of abstracting away from the evaluation of the correspondence relation by Faithfulness constraints is to exclude from consideration all but the "best" correspondence for each structure. In tableau (4), for example, *.b* a_2. and *.a* . have gratuitous MAX violations, but .b_1a_2. and .a_2. do not, so are more natural. We use this notion here. For each structure we can choose whichever correspondence relation leads to greatest harmony according to the constraint hierarchy. Call this correspondence the **most harmonic correspondence**. For example, such possibilities as gratuitous lack of correspondents for some structural elements in the input and output, or gratuitous multiple correspondents for some elements in the input and output, are excluded because they lead to excessive violations of Faithfulness constraints. We now have the notion of **structure harmony** in (5):

(5) **Relative Harmony of Structures**

Let I be an input form, O1,O2 some structures.

Let C1 be the correspondence relation $C1 \subseteq O1 \times I$ such that

$\forall C \subseteq O1 \times I$, $\langle I,O1,C1 \rangle \succ \langle I,O1,C \rangle$. Define C2 similarly.

$\prec^{(\text{structure})}$ is defined as an ordering on structures such that

$O1 \prec^{(\text{structure})} O2$ iff $\langle I,O1,C1 \rangle \prec \langle I,O2,C2 \rangle$

One structure O1 is less harmonic than another O2 if and only if the candidate containing O1 under the most harmonic correspondence is less harmonic than the candidate containing O2 under the most harmonic correspondence.

5.1.3 Formulating The Analogy

Now that a notion of relative harmony of structures has been properly derived from the harmony relation on candidates, the derivation/harmony analogy may now be expressed as in (6):

(6) **Derivation / Harmony Analogy**

Let D=P1, ..., Pn be a derivation.

Let H be a harmony scale in which Pn is optimal

D and H are analogous to the extent that, for i=1,...,n-1:

if Pi *is succeeded by* Pi+1, then also Pi *is less harmonic than* Pi+1

Of course, harmony relationships exist between many more pairs of structures than just those which also happen to be in the derivation, so the analogy between D and H is tested just for those specific structures that are in D, seeing whether the succession relationships will be matched by harmony relationships.

A formally more thorough-going analogy can be achieved if we take into account the analogy between input-output correspondences and derivational-history relations which specify how structures in a derivation correspond with the original structure (chapter 3). The fuller analogy in (7) obtains between derivational sequences and harmony scales where the modifications to the underlying structure are the same for structures in the derivational sequence and in the harmony scale, i.e. that derivational histories are always equal to the most harmonic correspondence relations.

(7) **Derivation / Harmony Analogy (advanced)**

Let D=P1, ..., Pn be a derivation; for each i, let ${}_1H_i$ be the derivational history relation for Pi in the derivation.

Let H be a harmony scale in which Pn is optimal; let C_i be the most harmonic correspondence between input and the possible output Pi, for each i.

D and H are analogous to the extent that, for i=1,...,n-1:

(i) ${}_1H_i = C_i$; (ii) if Pi *is succeeded by* Pi+1, then also Pi *is less harmonic than* Pi+1

A Duke of York gambit of deleting and re-inserting an element, for example, would always fail condition (7i) of this more thorough-going analogy since the equivalent MAX and DEP violations could never be the most harmonic correspondence (faithfully mapping the element is better). In typical cases (7i) is a reasonable demand, for, as shown in 3.3, derivational histories and input-output correspondences take the same intuitively natural forms to the extent that violated Faithfulness constraints are ranked as low as possible ("Constraint violation is minimal") and derivations are as short as possible (Economy of derivation), except perhaps in a few formally subtle cases.

5.2 The Extent of the Correlation

Having formulated the analogy between derivational succession and harmony incrementation in a formally defensible way, we now test the actual extent to which traversing along the steps of a derivational sequence to the end is consistent with traversing through increments in harmony up to the peak.

Moreton (1999) has proven the key property of Harmonic Ascent for constraint evaluation systems, according to which the output form given by a hierarchy of Markedness and Faithfulness constraints is either identical to the input or more harmonic than the input – as a result of satisfying high-ranking Markedness constraints. As far as comparison with derivations is concerned, this means that:

- if the output is identical to the input, then the analogous derivation is one with no steps at all, so the derivation-harmony correlation is vacuous;
- if the output is different and therefore more harmonic than the input, it follows that a derivation which takes a path from one to the other corresponds to an *overall* increase in harmony, ending on a more harmonic form than the one it starts on.

The question now is whether or not this increase is distributed over each one of the individual steps of the derivation.

5.2.1 Last Step as Increment

Whenever a derivational sequence and a harmony scale converge on the same surface form, the following result in (8) obtains: the last step in the derivational sequence always corresponds to a harmony increment, since the final form is also the optimal one on the harmony scale, whereas the penultimate form in the sequence is suboptimal on the harmony scale.

(8) **Last Step as Only Necessary Increment**

Let D=P1,...,Pn, be a derivation; let H be a corresponding harmony scale converging on the same surface form as D.

a. *In the derivation, each structure except the last is immediately succeeded by another.*

for i=1,...,n-1   Pi $\rightarrow$ Pi+1

b. *Each structure that happens to take part in the derivation apart from the last is suboptimal on the harmony scale.*

Assuming that D and H converge on the same surface form (Pn), then

for i=1,...,n-1   Pi $\prec$ Pn

c. *Derivational succession and harmony incrementation necessarily correlate in the last step, and this is the only necessary point of correlation.*

The relational statements in a. and b. match iff i=n-1:

$\text{Pn-1} \rightarrow \text{Pn}$ & $\text{Pn-1} \prec \text{Pn}$

The result follows on the assumption that a derivational sequence and a harmony scale converge on the same surface form. This assumption is not entirely trivial, on two counts. First, as was shown in chapter four, differences in output can be thrown up solely from the differences between rule interaction and constraint interaction, even when the rules and constraints themselves are strongly analogous. Second, as drawn attention to by Hammond (2000), it is possible to construct evaluation systems that have two or more optimal forms, with no constraint to discriminate between them. This contrasts with derivational sequences, since a sequence has precisely one final member. Scales thus depart from sequences in this essential respect. We consider *how* to restrict scales to a single optimal output in an appendix to this chapter.

The result in (8) opens up a difference between the last step, where the derivation-harmony correlation is guaranteed, and other steps where it is not guaranteed, it is now inevitable that the correlation between derivational succession and harmony incrementation is limited. While some derivations may be entirely consistent with harmony increments, the possibility of troughs, plateaus, or peaks, as illustrated in (9), remains.

(9) Derivation $\rightarrow$

					[P6]	
	P2	P3		P5		$\uparrow$ Harmony
/P1/			P4			

Such mismatches do indeed occur. This is illustrated from a simple example due to Prince and Smolensky (1993:206-207). Consider a rudimentary grammar which admits CV(C) syllables, and which delivers epenthesised forms for aberrant input sequences failing to comply with CV(C). Thus, given the input */V/*, the grammar will derive a syllable consisting of the *V* augmented with an epenthetic consonant to provide the necessary onset: *.cV.* . A rule-based system might achieve this by a syllable formation rule and an onset-consonant epenthesis rule, but an alternative OT grammar would have constraints ONSET (syllables have onsets), NOCOMPLEX (each syllable subconstituent contains just one segment), and PARSE (segments must be parsed into syllable structure). These constraints are undominated, but FILL (constituents must be filled by underlying material) is crucially ranked below them so that epenthetic positions may be admitted so as to comply with the requirements of syllable structure. The following table (10) cites the rule applications deriving .cV. from .V. from V (entered to the left of the forms) opposite the constraint violations of the same forms given by the OT grammar (to the right of the forms), presenting a "history" of constraint violations for structures that are found in the derivational sequence.

(10) **Constraint Violation History** (Prince and Smolensky 1993:207)

Stage	Rule	Representation	ONSET	PARSE	FILL
1		V		*	
2	Nucleus Formation	.V.	*		
3	Onset Epenthesis	.cV.			*

As anticipated, the last step corresponds to a harmony increment. The last step is .V.→.cV. and

.cV. is more harmonic, by satisfaction of ONSET at the expense only of FILL. However, the previous step V→.V. does not correspond to a harmony increment, since each registers one violation from an undominated constraint. Hence, a move from stage 1 to stage 2 along the derivational sequence constitutes a harmony plateau.

We could, of course, remedy this artificially by ranking the constraints arbitrarily. We could rank PARSE above ONSET, so that the ONSET-violating form .V. is more harmonic than the PARSE-violating form /V/ (though still not better than .cV.). Equally, however, we could rank the other way - ONSET above PARSE - so that the relation between stages 1 and 2 corresponds to a drop in harmony (though the relation between stages 2 and 3 still necessarily corresponds to a rise in harmony). In this case, then, the succession from stage 1 to stage 2 is *ambivalent* to harmony.

5.2.2 Postconditions and Restraints

We can further the analysis by considering the general properties of constraints that would be relevant to the two structures involved in a derivational step. A derivation is made up of steps containing two minimally-different representations, one of which is subject to the application of a rule and the other of which is the result of applying the rule. In an evaluation, representations are partitioned by each constraint according to how many violations they incur. So whether a derivational step corresponds to harmony incrementation or not depends on whether the constraint violations accruing to the second of the two structures are fewer, or belong to lower-ranked constraints, than those of the first.

Harnessing terminology due to Prince and Smolensky (1993:206), a constraint according to which the succeeding structure is more harmonic than the structure it succeeds we may call a **postcondition**, and a constraint according to which the succeeding structure is less harmonic than the structure it succeeds we may call a **restraint**, as in (11). A recurrent postcondition/restraint contrast is between a markedness constraint demanding some change to the representation versus a faithfulness constraint disfavouring the change, though the concept is general enough to take in other contrasts (the postcondition/restraint “contrast” may constitute a constraint *conflict* - but not necessarily, as we will show).

(11) rule

$P1 \rightarrow P2 \rightarrow \ldots \rightarrow Pi \rightarrow Pi{+}1 \ldots \rightarrow Pn$

*   √   ‘postcondition’

√   *   ‘restraint’

discriminating constraints

A constraint may express generalisations of a subtlety different form than the rule to which it is a postcondition, or restraint. The requirement that syllables have onsets represents a postcondition to the rule of onset formation. Yet, while the representation resulting from onset formation on a <C>.V.<C>.V.<C>.V. string produces .CV.CV.CV with no ONSET violations, exhaustive application of onset formation on .V.<CC>.V.V.<C>V gives .V<C>.CV.V.CV. which has fewer ONSET violations but still retains some. The difference arises because the ONSET constraint is a constraint on syllables whereas the onset formation rule is a rule about phonemes. As Roca (1994:145) observes, the principle of disallowing onsetless syllables is only satisfied by onset formation in the presence of suitable segmental material. In fact, many constraints formulated in Optimality Theory either require or are predicated over syllable structure and higher prosodic and metrical structure. Postconditions and restraints thus make more sense if applied to the *syllabification* - or for that matter, *prosodification* - of the output of each rule. This accords with the proposal in rule-based theory that syllabification re-applies to the output of each rule throughout the derivation (McCarthy 1979, Itô 1986). If we assume this, then each step Pi,Pi+1 in the sequence itself contains a mini-sequence containing the application of some rule plus rules assigning prosodic structures.

It is clear from the diagram (11) that postconditions are consistent with the analogy between derivational succession and harmony incrementation, while restraints are directly contrary to it. But the most highly ranked constraint is always decisive in optimality theory. So the analogy will hold to the extent that for each pair of representations Pi, Pi+1 there is some postcondition that dominates all restraints. Now constraint ranking is settled by discriminating between the optimal form and every other form, as in (12). Only at the last step will it always be the case that a postcondition dominates. Postcondition/restraint analysis thus recapitulates the result of the previous section 5.2.1.

(12) **Dominant Postcondition for Last Step**

Constraint Ranking Logic:

If P_{opt} is the surface form and P_{subopt} another form, then there must be a constraint C which rejects P_{subopt} in favour of P_{opt}, which dominates all constraints C_x which reject P_{opt} in favour of P_{subopt}. (Without C, P_{opt} will not be optimal.)

	C	>>	Cx
P_{opt}			*
P_{subopt}	*		

Corollary 1: Among those forms involved in a derivational sequence P1,...,Pn, where Pn is the surface form, then for i=1 to n-1, there must be a constraint Ci which rejects Pi in favour of Pn, which dominates all constraints Cx which reject Pn in favour of Pi.

Corollary 2: There must be a postcondition which rejects Pn-1 in favour of Pn, which dominates all restraints which reject Pn in favour of Pn-1. (From corollary 1, with i=n-1)

5.2.3 Typical Derivational Sequences

When we consider postconditions and restraints at *any* step in a derivational sequence, not just the first, there are two kinds of histories that postconditions may have. The first, represented in (13a), is where a postcondition is violated throughout an initial portion of the forms in the sequence P1,...,Pi, but satisfied by the remainder. The second, represented in (13b), is where a postcondition is satisfied by forms in an initial portion, but violated by a further form or forms, and satisfied by the remainder:

(13) **Unfed rules and Fed rules**

a. P1 → P2 → ... → Pi → Pi+1 ... → Pn

P1	P2	Pi	Pi+1	Pn	
*	*	*	✓	✓	‘postcondition:i+1’
✓	✓	✓	*	*	‘restraint:i+1’

b. P1 → ... → Pi → Pi+1 → Pi+2 ... → Pn

P1	Pi	Pi+1	Pi+2	Pn	
*	*	✓	✓	✓	‘postcondition:i+1’
✓	✓	*	✓	✓	‘postcondition:i+2’
✓	✓	✓	*	*	‘restraint:i+2’

Whereas (13a) reflects a rule whose structural description is met at the outset of a derivation, (13b) reflects a rule whose structural description is fed by another rule in the derivation. In (13a), succession corresponds to harmony incrementation, since a postcondition must dominate all restraints by comparing their violations for Pi and Pn, shown in (14a):

(14) **Ranking Arguments**

a. No feeding interaction in derivation

	postcondition:i+1	>>	restraint:i+1
Pi	*		✓
Pn	✓		*

b. Feeding interaction in derivation

	postcondition:i+1	>>	restraint:i+2
Pi	*		✓
Pn	✓		*
	postcondition:i+2	>>	restraint:i+2
Pi+1	*		✓
Pn	✓		*

For the feeding interaction, however, no ranking argument can be formulated between the two postconditions because they do not actually conflict: both are equally satisfied by the optimal form. This despite the fact that at the step (i,i+1), one (postcondition-1) is a postcondition and one (postcondition-2) is a restraint. The only ranking arguments that can be made are between these constraints and constraints which are violated by the surface form, as given in (14b). This includes the application of a “repair”-rule to the output of another rule, which is one use of the feeding interaction in rule theories. So, as Prince and Smolensky (1993:205ff) observe, the postcondition of the rule whose output is to be “repaired” (e.g. nucleus formation for a ‘V’) and the postcondition of the repair rule (e.g. onset epenthesis) do not conflict, since both are satisfied in the surface form.

It is also possible that a derivation may fail to converge on the harmony peak determined by the constraints which act as postconditions and restraints on the derivational steps. The path may overshoot or undershoot. Overshoot describes a case where there are more disparities between the final form in the derivation and the underlying form than between the optimal form and the underlying form, and undershoot describes a case where there are fewer disparities between the final form in the derivation and the underlying form than between the optimal form and the underlying form. These possibilities stand outside the assumption made in 5.2.1 that the derivational sequence and the harmony scale converge on the same surface form, but if we relax this assumption and examine the postconditions and restraints at each step of overshooting and undershooting derivations, we find in (15) that counterbleeding derivations positively correlate with the harmony scale, even though the path circumnavigates the most harmonic form and fails to converge on it.

(15) **Undershoot and Overshoot**

a. Undershoot: counterfeeding (feeding alternative)

$P1 \rightarrow P2$	$P1 \rightarrow P2 \rightarrow P3$	
* ✓	* ✓ ✓	‘postcondition-2’
✓ *	✓ * ✓	‘postcondition-3’
$P1 \approx P2\ (\prec P3)$	$P1 \approx P2 \prec P3$	

b. Overshoot: counterbleeding (bleeding alternative)

$P1 \rightarrow P2 \rightarrow P3$	$P1 \rightarrow P2'$	
* ✓ ✓	* ✓	‘postcondition-2’
* * ✓	* ✓	‘postcondition-3’
$P1 \prec P2 \prec P3\ (\prec P2')$	$P1 \prec P2'$	

In a counterfeeding interaction (15a), just as one rule applies but the other fails to apply afterwards, so also one postcondition is alleviated while another postcondition is left violated. By contrast, when one rule feeds the other, both postconditions are satisfied. In a counterbleeding interaction, portrayed in (15b), two constraint violations are alleviated over two steps, both matching with harmony increments. Although the counterbleeding derivation is consistent with harmony incrementation it does not lead to the harmony *maximum*, but rather skirts it. The output obtained by the bleeding derivation (P2') will be more harmonic than the end-point of the counterbleeding derivation (P3), because when one rule bleeds the other, removing the need for it to apply, both constraints are alleviated in one step, a more faithful alternative. Overshoot and undershoot offer a mixture of advantages and disadvantages empirically (see chapter four and chapter six): overshooting derivations provide the correct results in "overapplication", but precisely the wrong results in cases of "default" effects; undershooting derivations fail to derive "mutual interdependence" effects, though they allow "underapplication" effects to be described.

5.2.4 A Duke-of-York Derivation: Irreducible Harmony Drop

At steps caused by rules which feed, where there is no conflict between postconditions and restraints, there is no basis for a derivational step which *necessarily* leads to a *less* harmonic form. It remains to ask whether there are any such cases among derivational sequences and harmony scales which converge on the same surface form. This possibility is found in Duke-of-York derivations. In the Duke of York Gambit, some structural change A→B is followed in the derivation by the reverse change B→A. An illustrative example in (16) is from Nootka (Pullum 1976:94, from an unpublished paper by Campbell).

(16) ...ok

$\downarrow$ Labialisation $k \rightarrow k^w$ / o__

...ok^w

$\downarrow$ Delabialisation $k \rightarrow k^w$ / __#

...ok

Nootka has labialised and unlabialised dorsal stops. Labialisation is always removed word-finally, but dorsal stops are always labialised following an *o*. But in the overlapping context *o__#* dorsal stops are not labialised, so Delabialisation must be ordered after Labialisation to ensure this. Then, given an underlying form ending in *...ok#*, both rules apply in turn to leave the stop unlabialised at the end.

On a harmony scale, the form *...ok* will be optimal, and hence more harmonic than *...ok^w* . This is supported by the following tableau:

(17)

/...ok/, /...ok^w/	No final *k^w*	No *k* after *o*
☞...ok		*
...ok^w	*!	

As ever, the final step of the derivational sequence 〈 ...ok^w , ...ok 〉 corresponds to a harmony increment, but since the preceding step 〈 ...ok , ...ok^w 〉 is the inverse of the last step, it inevitably corresponds to a harmony drop.

In general, a derivational step leads to a less harmonic form whenever there is some restraint that dominates all postconditions at that step. Assuming that the derivational sequence and harmony scale converge on the same output, this happens when the surface form Pn violates one of the postconditions, so that some restraint must dominate that postcondition to guarantee the optimality of Pn.

(18) **Harmony Drop**

a.	P1 →	... →	Pi →	Pi+1	...	→ Pn	
			*	✓		*	‘postcondition’
			✓	*		✓	‘restraint’

b.

	restraint >>	postcondition
Pi+1	*	✓
Pn	✓	*

This means, as the derivation (18a) illustrates, that if there is a derivational step corresponding to a drop in harmony, it must be that the remainder of the derivation contains changes that reverse the effect at that step. A Duke of York gambit does this.

5.2.5 Summary

Unlike the other derivational patterns we have examined, a simple Duke of York gambit involves a step that necessarily contradicts harmony. Feeding interactions involve a step that is ambivalent to harmony, depending exactly on how the constraints are (arbitrarily) ranked. Other derivations – even derivations that overshoot the result given by the constraints that are postconditions and restraints for the steps of the derivation – correspond to increments in harmony.

5.3 Restricting Sequences

In 5.2.4, it was shown that where a step corresponds to a harmony drop, the change at that step is reversed later in the derivation. If we excluded derivations with such reversals we would eliminate all derivational steps that correspond to a harmony drop, achieving a closer match between derivation and harmony. We begin by considering a ban on Duke-of-York derivations, but the derivation-harmony mismatch goes deeper: we soon show that sequences themselves are mathematically different from scales.

5.3.1 Excluding Duke-of-York Derivations

Pullum (1976) observed that generative phonologists have often expressed misgivings, somewhat inchoately, about Duke-of-York derivations and attempted to avoid them. Reviewing this phenomenon in chapter three, we argued that Duke-of-York derivations are unexplanatory and generally unsupported empirically. They fail to explain the similarity of surface forms to their underlying forms, and in crucial cases of languages with both vowel deletion and insertion, where deletion-insertion derivations would be detectable by impoverishment of the vowel inventory, they fail to occur.

Excluding Duke-of-York derivations would eliminate a class of derivations which have a step that goes down the harmony scale instead of up. And it would force us to re-analyse putative cases - labialisation in Nootka (5.2.3.) need not apply word-finally where it would be reversed if it is confined to dorsals in syllable onset, or to prevocalic dorsals whose release phase is more amenable to carrying labialisation audibly. However, as well as excluding some Duke-of-York gambits that do *not* have a step that goes down the harmony scale (see 5.3.3 below), the move would fail to exclude other derivational sequences that do. In simple Duke-of-York derivations, a structure is repeated in two steps when one structural change is immediately reversed by the opposite change, but a structure could conceivably be repeated under different conditions, as the hypothetical derivations in (19) demonstrate.

(19)

a. pai

pii	Raising *a*
pi:	Vowel Deletion in presence of identical vowel
pai	Diphthongisation

b. prar

parr	Liquid/Vowel Metathesis
parar	*a*-Epenthesis
prar	Vowel Deletion between Stop and Liquid

Neither of the hypothetical derivations in (19) employs a Duke-of-York gambit. The first ends in a fissure of one vowel into two, re-creating a diphthong of which one half was deleted. The second has contrary structural changes, insertion and deletion of *a*, but these apply in different positions and a third rule, a metathesis, completes the re-creation of a previous structure. A structure might be repeated by some even more convoluted set of changes. Excluding Duke-of-York derivations only deals with the more obvious cases.

Yet all sequences that have a repeated member, Duke of York gambit or not, necessarily contain a step which corresponds to a drop in harmony, conforming to the scheme given in (18). In the harmony evaluation corresponding to the derivation *prar* → *parr* → *parar* → *prar* (19b), the constraint requiring metathesis of *a* and *r* (a postcondition at the first step) is violated by the final form, so must be dominated by a conflicting constraint that is satisfied in the final form. The constraint requiring metathesis is violated when it would create *rr*, so is dominated by a constraint against *rr*. Then the step *prar* → *parr* is a drop in harmony because it creates *rr*. For all repetition after two or more steps, the situation is as in (20) (illustrated for repetition after three steps for concreteness):

(20) **Structure Repeated in Sequence**

P1 → ... →	Pi →	Pi+1→	Pi+2→	Pi	... →	Pn	
	*	✓	✓	*		*	postcondition:i+1
	✓	*	✓	✓		✓	postcondition:i+2
	✓	✓	*	✓		✓	postcondition:i+3

At the step which moves away from the structure which is eventually repeated, any postcondition is eventually violated again. Assuming the rule which changes the repeated structure does not apply second time round (nor does any other rule change any part of its structural description, which would create an overshoot), then the constraint violation will persist among subsequent members of the sequence, including the final surface form. This leads to a ranking argument: some other conflicting constraint must be satisfied at the expense of postcondition:i+1. A postcondition at the second step after the structure to be repeated is such a constraint, being both a restraint at the first step and satisfied by the surface form.

(21) **Ranking Argument** for (20)

	postcondition:i+2	>>	postcondition:i+1
Pi+1	*		✓
Pn	✓		*

This takes care of repetition after two or more steps. Repeating a structure in one step can only happen with a rule capable of producing an output identical to its input. Out of the usual set of operations (see 3.1), the only possibility is a rule of metathesis that sometimes interchanges identical elements, e.g. $pr_i ar_j \rightarrow pr_j ar_i$. This would correspond to a drop in harmony since the input and output would violate constraints identically except that the output has an additional violation of a faithfulness constraint (LINEARITY, in the case of metathesis; and if it happened twice in a derivation, the second time would correct the LINEARITY violation but the first time would still correspond to a harmony drop).

5.3.2 Sequences are not Ordered Sets

It is possible for a sequence to contain a member that is repeated in the sequence, but this does not happen in harmony scales, for each entity in a scale has its own place in the scale. This sets sequences apart from scales mathematically.

An *ordering* is a relation that is irreflexive, asymmetric and transitive, whose definitions are given below:

(22) Let A be a set. Let R be a ordering relation in A. The following are true of R:

a. **Irreflexivity**: $\forall a \in A$, it is not the case that aRa.

No element is ordered before itself.

b. **Asymmetry**: $\forall a_1, a_2 \in A$, if $a_1 R a_2$ then it is not the case that $a_2 R a_1$.

A pair of elements can only be ordered one way.

c. **Transitivity**: $\forall a_1, a_2, a_3 \in A$, if $a_1 R a_2$ and $a_2 R a_3$ then $a_1 R a_3$.

The order of two elements is settled if there is an intermediary ordered between.

The relation "less harmonic than" is an ordering: (a) no representation is less harmonic than itself (irreflexivity), since it only has one set of constraint violations; (b) one representation cannot be both less harmonic and more harmonic than another (asymmetry), since the highest-ranked constraint on which they differ settles one way or the other; (c) if one representation is less harmonic than another, it is less harmonic than representations that the other is less harmonic than (transitivity), because the one representation has more serious violations than the other, and these are more serious than still less costly representations.[1]

It would likewise seem initially plausible to suggest that derivational sequences are ordered sets, albeit smaller, finite ones. One can recognise on sequences a relation of immediate succession e.g. ‘P2 immediately succeeds P1’ and from it a general relation of succession e.g. ‘P5 succeeds P1’. These properties fail, however, in sequences that have a repeated member, because a repeated member of the sequence succeeds itself. For example, in two steps a Duke-of-York gambit gives A→B→A giving rise to the relations for each step ‘B succeeds A’ and ‘A succeeds B’ (failing asymmetry) and then overall ‘A succeeds A’ (failing irreflexivity). This shows that sequences are not the same as ordered sets.

Sequences still make formal sense even if they are not ordered sets. Members of a sequence have a “place” in the sequence; repetition is when a member of the sequence has more than one place in the sequence. The places in the sequence are ordered, even if the members are not. It is often convenient to think of the “places” as numbers - as we do when we cite a derivational sequence as P1,P2,P3,...,Pn . This leads us to a definition of sequences, as given in (23).

---

[1] Other essential properties of the “less harmonic than” relation include that: (i) it has a *greatest* element (there is at least one optimal element); (ii) it defines a *partition* into *equivalence classes*, whose members are characterised by the same degree of violation and which share the same ordering relations to members of the other classes; (iii) it is not *connected*, since there are pairs of structures that are not ordered by harmony one way or the other (i.e. any pair in the same equivalence class).

(23) **Definition: Sequences**

a. A **sequence** is a triple ⟨M,P,A⟩ where M is a set of **members**, P is a set of **places**, which is well-ordered with a least element (*e.g., the set of natural numbers*), and A is a function from P onto M.

Each place in a sequence starting from the first is assigned a member.

b. A sequence is **ordered** if A is a one-to-one correspondence.

If each member of the sequence occurs only once, members will be ordered incidentally along with their respective places.

c. A sequence is **finite** if the domain of A in P is finite.

If only a finite number of places are assigned members, the sequence terminates.

In a derivational sequence where there is no repeated member, the members and places are then in one-to-one correspondence, so that particular derivation at least is ordered. Derivational sequences are always finite, because they terminate after a finite number of places (i.e. stages), although derivations are not bounded by any particular limit (e.g. “10”, or “99”). The finiteness of derivations is significant in connection with derivations that contain a repeated member, for it means that rules cannot re-create a structure again and again indefinitely. Unconstrained re-application of rules must be prevented either by strict ordering or by the constraint that rules cannot apply in interrupted sequence (Ringen 1976), and by the regularity constraint that rules cannot re-apply to the new configuration in their own output (Johnson 1972, Karttunen 1993).

We return to the differences between sequences and scales in an appendix to this chapter. For while sequences exceed the orderedness of scales in general, scales exceed the connectedness of sequences in general.

5.3.3 Derivational Economy

Following (23b), derivational sequences with repeated members may be excluded directly by the condition of orderability given in (24a), which admits only sequences whose members have just one place in the sequence. Alternatively, the economy condition (24b) which minimises the length of derivations, also excludes these sequences. Economy of derivation has been explored as a principle of Minimalist linguistic theory (Chomsky 1995, Calabrese 1995).

(24) Conditions that exclude sequences with repeated members

a. **Orderability**: Derivational sequences are ordered.

b. **Economy**: Derivational sequences are of the minimum length possible to derive their final member from their initial member.

If a structure is repeated in a derivation, P1, ... Pi, ... Pi, ... Pn, then the derivation is replaceable with a shorter one which lacks the portion of the sequence Pi, ..., Pi that comes between the repeated tokens. This excludes not only Duke-of-York gambits (which exceed the minimum length possible by precisely two steps), but any other collection of rules which re-create a structure found earlier in the sequence. In fact, the economy condition is stronger than the orderability condition. The economy condition, but not the orderability condition, would select the derivation in (25b) over the one in (25a):

(25) a. *{wati → waɾi → waɾ → wat}

b. wati → wat

In the eliminated derivation (25a), the final step re-creates a /t/ removed at an earlier step. There is no repeated structure, but it is not the shortest possible derivation. The economy condition also inveighs against some other derivations (e.g. $\varnothing \rightarrow +F \rightarrow -F$ exceeds the shorter $\varnothing \rightarrow -F$).

Duke-of-York derivations, and derivations in which a structure is repeated, typify a certain family of derivations. There is some configuration within a structure that is altered and subsequently reconstructed (by Duke-of-York gambit, or some other way) in the course of a derivation (possibly with other rules making other changes). This is always technically uneconomical. Furthermore, derivational steps that necessarily correspond to a harmony drop always fall within this family: for as found in 5.2.3, any derivational step that necessarily corresponds to a harmony drop must be reversed in a subsequent part of the sequence (assuming that the sequence converges to the same surface form as the harmony scale). Since the economy condition (24b) excludes this family, it follows that it also eliminates all derivations that necessarily contradict harmony (feeding interactions can be made to contradict harmony by an arbitrary ranking of the constraints, but do not *necessarily* contradict harmony). This gives us (26):

(26) **Derivation/Harmony Correlation under Derivational Economy**

Economical derivations that converge to the same output as a harmony scale do not flatly contradict the harmony scale at any step.

The condition that derivations and harmony scales converge to the same output makes this a very basic kernel of patterns, of course, excluding overshoot and undershoot but including bleeding and feeding effects (as shown in 5.2.3), and mutually-contrary processes *if* they are prevented from creating an uneconomical Duke-of-York derivation. It also excludes the more complex Duke-of-York derivations where the initial change feeds an intervening rule before the reverse change is made, which McCarthy (2003) has argued are unattested.

However, although derivational steps that necessarily correspond to a harmony drop always fall within the family of uneconomical derivations that alter and re-construct part of a structure, there are some within this family that do not contradict harmony scales. Some Duke-of-York derivations are simply a series of feeding interactions, as Pullum (1976) discovered. He offered the following hypothetical data and rules:

(27) Pullum (1976:89-90)

káti	'wallaby'
katínlu	'wallabies'/wallaby-PL
katenlóma	'by wallabies'/wallaby-PL-ERG

A. Penultimate vowels are stressed.

B. High vowels become mid in unstressed, closed syllables.

C. Final *n* deletes after mid vowels.

D. Final mid vowels become high.

The forms can be derived as follows:

(28)

a. /katin+lo+ma/ $\xrightarrow{A}$ katin ló ma $\xrightarrow{B}$ [katenlóma]

b. /katin+lo/ $\xrightarrow{A}$ katín lo $\xrightarrow{D}$ [katínlu]

		A		B		C		D	
c. /katin/		→	kátin	→	káten	→	káte	→	[káti]

The last derivation (28c) uses a Duke-of-York gambit in regard to the height of the second vowel. The vowel begins high, is lowered by rule B, and is then raised again by rule D. There is no repetition of any structure, because rule C intervenes between the gambit-rules B and D, so the final form is not the same as the form prior to B. Furthermore, the rules apply simply when their structural description is met, each feeding the next. It is even the case that all the rules are transparent: no rule alters the structural description of any previous rule. This Duke-of-York gambit may be replicated by a harmony evaluation:

(29) Constraints

A Stress is penultimate.

B No high vowels in unstressed, closed syllables.

C No word-final *n* after mid vowels.

D Word-final vowels are high.

/katin/	A	B	C	D	IDENT([high])	MAX(C)
katin	*!					
kátin		*!				
káten			*!		*	
káte				*!	*	*
☞káti						*

Even though the derivation (28c) does not contradict harmony as tableau (29) shows, it would be excluded by the economy condition. Of course, it is not positively consistent with harmony either. Comparison to the tableau shows that the derivational steps prior to the last correspond to harmony plateaus (unless some arbitrary ranking of the constraints is made), in accordance with the pattern for feeding interactions observed in 5.2.2. This shows that the possibility of feeding initiates open-endedly complex derivations that provide increasingly serious mismatches with harmony:

(30)

Seriousness of Mismatch	Derivations
Ambivalence to harmony	Feeding derivations – including those uneconomical derivations where the rule recreating structure changed by an earlier rule is fed by an intervening rule
Flat contradiction of harmony	All other uneconomical derivations in which structure is re-created after being altered – including all unorderable derivations.

Derivations partly similar to Pullum's example occur in the Slavic languages Slovak (Rubach 1993:267) and Polish (Rubach 1984:101), where Depalatalisation undoes the effect of Palatalisation when Depalatalisation is fed by the intervening rule of yer deletion. In Slovak, the derivation of *vodný* 'watery', adjective from *voda* 'water', runs as follows:

(31) 1 vod+En+ý

2 vod´+En+ý Palatalisation

3 vod´+n+ý Yer Deletion

4 vod+n+ý Depalatalisation

The latent "yer" vowel, in this case a front vowel capable of palatalising the root-final consonant, does not vocalise in this context, leaving a consonant cluster. This removes the conditioning environment for palatalisation - an overapplication effect. Because of yer deletion, the *d´* is now in the preconsonantal environment of Depalatalisation (as opposed to the prevocalic environment of Palatalisation), a feeding effect. As sketched in (32), this derivation too may be replicated on a tableau:

(32)

/vod+En+ý/	Palatal (__FrontVowel)	No Vocalisation	No Palatal (__CoronalCons)
vodený	*!	*	
vod´ený		*!	
vod´ný			*!
☞vodný			

The Slavic pattern has in common with Pullum's hypothetical example the feature that the reversal rule is fed by an intervening rule. As tableaux (29) and (32) indicate, this corresponds to a case where constraints do not conflict since they trigger the contrary processes under disjoint conditions. This differs from examples of Duke-of-York derivations with intervening rules studied by McCarthy (2003). In the examples which McCarthy inveighs against, the intervening rule does not feed the reversal rule, but rather the original rule feeds both the intervening rule and the reversal rule. Those cannot be replicated by constraint interaction.

The result is that Duke-of-York derivations that are consistent with harmony have empirical currency, as in Slavic Depalatalisation. Other Duke of York derivations (i.e. all those which flatly contradict harmony) are unattested, a claim supported both by the universal absence of vowel deletion-and-insertion derivations (see 3.4.4, 4.4.2), and by the arguments in McCarthy (2003) against other putative Duke-of-York derivations with intervening rules. Hence, economy of derivation is just slightly too strong a condition, and instead it is consistency with harmony that emerges as the property that matches the empirical record.

5.4 Conclusion

The analogy between derivational sequences and harmony scales has real but limited currency. The succession-incrementation correlation holds solidly for the last step of a derivation that converges with a harmony maximum, but recedes as sequences and scales themselves recede from the mathematical properties held by the other. For sequences and scales are mathematically different, as demonstrated in 5.3.2 and in the appendix to this chapter.

Some derivations (those with feeding interactions) contain a derivational step that is ambivalent to harmony; others (most Duke of York derivations) flatly contradict harmony. Derivational steps which flatly contradict harmony are eliminated by the condition of derivational economy. However, this would exclude other derivations which do not flatly contradict harmony, and hence would exclude the palatalisation patterns in Slavic languages. Whereas the economy condition is too strong, it appears that derivations which are consistent with harmony should be admissible, but derivations that contradict harmony should not.

Appendix: Scales Are Not Connected Sets

Sequences differ mathematically from scales in that sequences are not ordered sets (5.3.2), but also in that harmony scales are not *connected sets*: constraints do not necessarily discriminate between every last pair of representations. By contrast, the property of succession differentiates between every pair of members in a sequence. Hammond (2000) points out that it is possible to construct evaluation systems that select several outputs as optimal. There could be two candidates *a*,*b* that are maximally harmonic with respect to the available constraints, and no constraint that further discriminates between *a* and *b*. Sequences intrinsically have just one final member, so always provide one unique output (caveat: assuming rules are "obligatory" rather than "optional"). We pursued our comparison of sequences and scales that have the same final/optimal output, but harmony scales are only comparable with sequences to the extent that they have a unique optimal form. This being so, we consider how to foreclose the possibility of multiple optimal forms.

Successive constraints in a hierarchy pare down the candidate set's most harmonic forms, but ending up with just one depends on "enough" of the "right" constraints. Individual phonologists can always propose constraints to get the single output they want in a particular case, but if we consider the conditions under which multiple outputs could logically arise, we can see what is required to deliver only one optimal output in general. We now show that there are two definitive possibilities for generating multiple outputs, in evaluation systems that employ markedness and faithfulness constraints.

Suppose two candidates *a* and *b* are both optimal. Either they have identical output structures under different correspondence relations, or they have different output structures - if they have identical output structures this presents no problem - they still give a unique output, however subtle and unusual this might be. Suppose instead, then, that *a* and *b* have different structures. Either (i) they have different faithfulness-constraint violations or (ii) they have the same faithfulness violations.

Suppose first they have (i) different faithfulness violations. On (at least) one constraint F, *b* is less faithful than *a* (say). If *b* is still equally optimal, it must be more harmonic than *a* according to another constraint C ranked equally with F:

(A1)

	F	C
a		*
b	*	

In this scenario, however, the constraints F and C do at least discriminate between the possible outputs. A unique output may be obtained simply by ranking them. F>>C selects *a* uniquely; C>>F selects *b* uniquely. This can be resolved in one of two obvious ways. Either (i) we require that rank order of constraints is total, so we remove the possibility that two outputs can emerge from the equal ranking of two normally conflicting constraints, or (ii) we state as an added axiom the requirement that only one output be accepted, and this axiom will force constraints that favour alternative outputs to be ranked one way or the other.

Suppose next that two outputs have the same faithfulness violations - case (ii). Since *a* and *b* are different structures, there must be (at least) one faithfulness constraint that is violated at two different positions - two different disparities of the same type on the same tier. Schematically, we can present *a* and *b* as containing the following structures, disparities of the form *p*→*q* in contexts *x_y_z*:

(A2)

input:	*xpypz*
a:	*xpyqz*
b:	*xqypz*

If the disparity is motivated in both contexts *x_y* and *y_z*, why are both not carried through?: why is *xqyqz* not optimal? It may be that one change is enough, for example if *p*→*q* is an insertion of an epenthetic vowel into a cluster of consonants CCC→CVCC, CCC→CCVC, then one epenthesis may be enough to create an acceptable syllable structure of consonants and vowels. But if there is still motivation to instantiate a disparity *p*→*q* in two places, there must be some dominant constraint blocking *xqyqz*. This could conceivably be a constraint **...q...q...* creating a dissimilation effect, e.g. ruling out successive identical vowels or consonants. Or if the *p*→*q* disparity is in fact one of deletion *p*→∅, e.g. vowel deletion CVCVC→CVCC,CCVC,CCC, there could be limitations **xyz* on the resulting strings such as unacceptable consonant clusters.

If we wish to rule out the possibility that multiple outputs arise by violation of one Faithfulness constraint at different positions, we must distinguish between *a* and *b* by constraints which distinguish different positions in a structure. Such constraints are necesssary, for example to place epenthetic vowels correctly into C*i*CC rather than *CC*i*C in Yokuts (McCarthy 2002:58), recapturing what in rule theory would be achieved by indexing the structural change ∅→*i* to the correct place in the structural description C_CC. We might, for example, adopt a constraint which awards a violation for every segment separating a vowel from the left edge of the word. This follows the NOINTERVENING constraint family (see 2.4.2). In general, process placement can be achieved in one of two ways. Either (i) we assume that all constraints that would ever be needed to distinguish alternative sites are present in all grammars, or (ii) we state as an added axiom the requirement that only one output be accepted, and this axiom will force the construction of constraints as required that eliminate all but one output.

The available options do not explain why there would be only one output from a grammar. An axiom would be essentially stipulative. Alternatively, the necessary constraints must be *both* already present *and* totally ranked to get the right results. Perhaps instead the maintenance of a unique output where multiple outputs would be possible is explained by functional considerations of simplicity of expression and communication. If these considerations apply loosely, we would predict that the specific kinds of variability in grammatical forms predicted by evaluation systems – both optional processes and variable placement of a process in structure – will be frequent in language variation.