Cross-level interactions in Harmonic Serialism

John J. McCarthy, University of Massachusetts - Amherst
Joe Pater, University of Massachusetts - Amherst
Kathryn Pruitt, Arizona State University
Cross-level interactions in Harmonic Serialism

John J. McCarthy, Joe Pater, & Kathryn Pruitt
University of Massachusetts Amherst

1 Introduction

Phonological generalizations must often be stated in ways that refer to more than one level of the prosodic hierarchy. For example, pre-Classical Latin words of the form CVCV: shorten the final long vowel to make a better trochaic foot: /amoː/ → (ˈamo) ‘I love’. This generalization refers to two levels of the prosodic hierarchy, foot and mora. We will refer to generalizations like this as cross-level interactions (CLIs).

In Kager (1999: 188-189), McCarthy (2002: 146-149), Mester (1994: 15), Pater (2000: 248-250), and Prince & Smolensky (1993/2004: 33-36, 148-149), CLIs are presented as crucial evidence for parallelism and against serialism in Optimality Theory (OT). The details of these arguments for parallelism will be discussed later, but for now it is sufficient to take note of what they have in common: they all involve CLIs that seem to require changing structure at two different levels of the prosodic hierarchy simultaneously. For instance, getting from /amoː/ to (ˈamo) in Latin seems to require building a foot and shortening a vowel in parallel. As Kager (1999: 189) puts it, “Parallelism establishes complete communication between all levels of the prosodic hierarchy. Consequently, changes at lower prosodic levels have effects at higher levels, and vice versa.” A serial analysis of the Latin facts has to shorten the vowel and build the foot in separate steps, but shortening before foot-building is unmotivated, and foot-building before shortening requires an intermediate stage (ˈamoː) with an ill-formed foot.

In this chapter, we review these arguments against serialism and show that they are not in fact arguments against at least one serial theory, Harmonic Serialism. Two ideas are key. One is violation of the surface-true: there are rankings in HS where a derivation introduces and later eliminates violations of a surface-unviolated constraint. The foot-structure constraint violated by Latin (ˈamoː) is an example. The other key idea is full availability of structural operations: in HS, structure is not built in a strictly bottom-up fashion, and conditions arising at later stages of the derivation can demand revision of structure inherited from earlier stages. Vowel shortening in the mapping (ˈamoː) → (ˈamo) is an example of such a revision.

To understand why CLIs have been seen as evidence for parallelism, it is necessary to understand why they are problematic in rule+constraint serial theories. In sections 2 and 3, we explain the problem that rule+constraint theories have with CLIs and how this problem is solved in parallel OT. We then go on in section 4 to argue that constraint violability, rather than parallelism, is essential to analyzing CLIs, leading to the conclusion that they are not problematic in HS. This is followed by discussion and further exemplification of violation of the surface-true in section 5 and full availability of structural operations in section 6. We conclude in section 7 with a demonstration that language typology puts limits on what kinds of CLIs could and could not be analyzed in HS.

2 Cross-level interactions in rule+constraint theories

It is helpful to approach the analysis of CLIs in HS by way of a historical antecedent: the effort to incorporate constraints into rule-based phonology. As we will show, CLIs do present serious issues for a serial theory, but only if it operates with inviolable constraints.

One way that constraints interact with rules is by blocking their application. Blocking by constraints was introduced by Kisseberth (1970):
(1) Blocking by a derivational constraint in Kisseberth (1970: 305)

A rule applies just in case

(a) its structural description is satisfied by the input string, and
(b) the output string would not be in violation of the derivational constraint.

Similarly, Liberman & Prince (1977) propose the following condition on the interaction of well-formedness conditions and rules:

(2) Blocking by constraints in Liberman & Prince (1977: 290)

No rule may apply so as to produce an ill-formed representation.

As an example of blocking, we can take the role of constraints in metrical theories of word stress that use both rules and constraints (Liberman & Prince (1977); see Hayes (1995), Kager (1995), and Kenstowicz (1994a) for further developments). In these theories, the parsing of syllables into feet is exhaustive, up to the limits imposed by the constraints that are active in the language. One CLI in this domain is the effect of the foot minimality constraint FOOT-BINARITY, which prohibits degenerate feet consisting of a single light syllable (Broselow 1982, Hayes 1995, McCarthy and Prince 1986/1996, Prince 1980, and much other work). Exhaustive parsing and the prohibition on degenerate feet come into conflict whenever odd-parity sequences of light syllables need to be parsed. Often, the leftover syllable remains unfooted, as in Hayes’s (1995: 205-208) analysis of Hixkaryana (Derbyshire 1979). In this language, iambic feet are assigned from left to right, and final syllables are extrametrical (indicated by angled brackets). (The effects of a later rule of iambic lengthening are not shown in (3) but will be discussed shortly.)

(3) Hixkaryana foot structure (after Hayes 1995: 207)

a. Even-parity sequence of light syllables before final
   (neˈmo)(koˈto)<no> ‘it fell’
   (ˈtoh)(kuˈrʲe)(hoˈna)(haˈja)<ka> ‘finally to Tohkurye’

b. Odd-parity sequence of light syllables before final
   (aˈʧo)wo<wo> ‘wind’
   *(aˈʧo)(ˈwo)<wo>
   *(ˈtoh)(kuˈrʲe)ho<na> ‘to Tohkurye’
   *(ˈtoh)(kuˈrʲe)(ˈho)<na>

Ungrammatical *(aˈʧo)(ˈwo)<wo> contains the degenerate foot (ˈwo). This is ruled out by FOOT-BINARITY, and the less complete foot parsing of (aˈʧo)wo<wo> is preferred instead. This is an example of the blocking mode of interaction between constraints and rules: a rule that attempts to parse exhaustively is blocked when the result would contain a degenerate foot.

Besides blocking, there is a second mode of interaction between constraints and rules, triggering. Triggering was first recognized by Sommerstein (1974), who introduced the idea of rules that are “positively motivated” by constraints:

(4) Triggering by constraints in Sommerstein (1974: 75)

A P-rule R is positively motivated with respect to a phonotactic constraint C just in case the input to R contains a matrix or matrices violating C AND the set of violations of C found in the output of R is null or is a proper subset of the set of such violations in the input to R.

According to Sommerstein’s convention on rule application, rules apply only when they are positively motivated according to the definition in (4), that is, when they remove a violation of a constraint. In short, the constraint violation triggers the application of the rule.

In triggering mode, FOOT-BINARITY could be enforced by lengthening a short vowel.
Assuming a moraic representation of length, the vowel lengthening rule can be stated as in (5):

(5) Vowel lengthening
   Insert a mora

Rules of this generality are typical in rule + constraint theories like the ones we are discussing. Under the principle in (4), the vowel lengthening rule in (5) will only be applicable when some phonotactic constraint triggers it.

An example of triggering can be drawn from Hayes’s (1995: 110-113, 208) analysis of what he calls the “unstressable word syndrome” in Hixkaryana. Words consisting of just two light syllables, such as /kana/ ‘fish’, offer up only one light syllable for footing, because the final syllable is extrametrical. In this case, FOOT-BINARITY triggers an application of the rule of vowel lengthening:

(6) Hixkaryana vowel lengthening triggered by FOOT-BINARITY

/kana/ → ⟨'ka⟩<na> → ⟨'ka⟩<na>

This example immediately raises the fundamental problem of formalizing rule + constraint theories. In (6), the constraint violation was introduced not by the underlying representation, as in Sommerstein’s examples, but rather by the prior application of the footing rule. If constraints are inviolable filters on the output of rules, the creation of the degenerate foot ⟨'ka⟩ should have been blocked. In some sense, this temporary violation of FOOT-BINARITY could be said to be licensed by the presence of the lengthening “repair”, but in a theory with serially ordered rules and inviolable constraints, it is hard to see how that explanation could be formalized.

One approach to the problem of how repairs license apparent constraint violations is to relax serial ordering, so as to allow the repair to apply in parallel with the rule that would otherwise create a constraint violation. This is the solution proposed by Mester (1994) for the Latin problem mentioned in the introduction. In a process called “iambic shortening”, Plautine Latin avoids the degenerate foot *⟨a⟩<mo⟩ by revoking the extrametricality of the final syllable and bringing it into the foot. The result would be an illicit light-heavy trochee *⟨amo⟩, so Latin also shortens the final vowel, yielding ⟨'amo⟩. In his analysis, Mester (1994: 15) conflates the building of the disyllabic foot and the shortening of the vowel into a single derivational step consisting of “a structure-changing imposition of a foot … on otherwise unfootable words. A structure-changing imposition of a foot is one which, simultaneous with constituent formation, invokes the designated repair strategy of the language (for Latin, REMOVE-µ …) to achieve a well-formed result.” By having the two rules apply simultaneously, the derivation goes directly from unfooted /amoː/ to footed and shortened ⟨'amo⟩, without intermediate stages that violate FOOT-BINARITY or the constraint against light-heavy trochees. Similar ideas are explored in Myers (1991) and Paradis (1988a, b).

While this locally solves the problem of temporary violation, as far as we know, this “repair strategy” approach has never been fully formalized. Hixkaryana illustrates the challenges that any theory of constraints, rules, and repairs will have to face. On the one hand, FOOT-BINARITY triggers the lengthening repair in /kana/ → ⟨'ka⟩<na⟩. But it also blocks footing of the penult in (3b) ⟨a⟩<fo⟩<wo⟩<wo⟩. Why isn’t the lengthening repair available in this case, yielding *⟨a⟩<fo⟩⟨'wo⟩<wo⟩? A clue is provided by Hayes’s identification of /kana/ → ⟨'ka⟩<na⟩ as an example of the unstressable word syndrome: evidently the lengthening repair is available only when an entire word would otherwise go footless. (Mester’s reference to “otherwise unfootable words” says practically the same thing about Latin.) But this restriction on when the repair is available goes unexpressed in the formal theory.

To make matters even worse, the lengthening repair is available in all words — not just those that would otherwise be footless — when the trigger is not FOOT-BINARITY but rather the
constraint SWP (for Stress to Weight Principle), which prefers light-heavy iambs to light-light ones. As a result of this process of iambic lengthening, as it is known, the actual surface forms of the examples in (3) have lengthening of stressed open syllables, as shown in (7):

(7) Result of iambic lengthening in Hixkaryana

Even-parity sequence of light syllables before final
(nem'no:)(ko'to:) < no >
(t'oh)(ku'ɾ'e):(ho'na):(ha'ɾa:) < ka >

b. Odd-parity sequence of light syllables before final
(a'lfo):wo < wo >
(t'oh)(ku'ɾ'e):ho < na >

Extant theories of rule/constraint interaction are unable to describe this state of affairs, where it appears that one constraint (SWP) is able to trigger a given repair (lengthening) but the other (FOOT-BINARITY) is not always able to do so (i.e., *(a'lfo)('wo:)<wo:>).

This is not an isolated example: many conspiracies involve complex interactions between triggering and blocking, which raise exactly the same questions. See, for example, the discussion of the formal challenges raised by Kisseberth’s (1970) Yawelmani conspiracy in Kiparsky (1973), McCarthy (2002, 2008), and Sommerstein (1977: 197-198).

Besides relaxing serialism, the other approach to the problem of violation and repair is to relax constraint inviolability. Something like this is suggested by Kisseberth (1970) in reference to an analysis of Tunica. In Tunica, some rules are blocked by a constraint against adjacent stressed syllables, while two other rules create these sequences, only to have them eliminated by later rules, thus raising for the first time the problem we are focusing on here. As a first step toward a solution Kisseberth offers a distinction between two kinds of rules, those that are relatively obligatory and those that are absolutely obligatory:

(8) Obligatoriness in Kisseberth (1970: 305)

A rule that is absolutely obligatory applies if its structural description is satisfied; it does not have to meet any added restrictions. A relatively obligatory rule is a rule which applies only if its structural description is satisfied and its output would not violate a given derivational constraint.

This idea will not help with blocking and repair by the constraint FOOT-BINARITY in Hixkaryana, however. The problem is that the same rule, foot formation, acts as absolutely obligatory with respect to FOOT-BINARITY in /kana/ → *(kə)nə → *(ka:)nə, but as only relatively obligatory with respect to FOOT-BINARITY in the blocked derivation /aʧo:wowo/ → (a'lfo)wo < wo > → blocked *(a'lfo)'wo < wo > → *(a'lfo)('wo:)<wo: >.

The literature reviewed in this section was largely successful in identifying the ways in which phonological processes can interact with well-formedness constraints. Formalization proved problematic, however — no rule + constraint theory was able to account for the range of blocking and triggering effects that are found cross-linguistically, as illustrated by Hixkaryana.

3 Cross-level interactions in parallel OT

It is clear from the discussion in section 2 that rule + constraint theories have real problems with Hixkaryana’s CLI. The worst of these problems are not shared by the parallel OT (P-OT) analysis of Hixkaryana in Kager (1999: section 4.3), however.

Parallelism allows GEN to produce candidates that show the effects of more than one operation simultaneously. Thus, GEN(/kana/) includes (kə:nə, which shows the simultaneous effects of both the footing and the lengthening operations. This allows it to beat *(kə)nə, with its violation of FOOT-BINARITY, *(kə:nə), which violates NON-FINALITY, and footless kana,
which violates the constraint requiring every prosodic word to contain a main-stressed head foot, \textsc{Headedness(PrWd)}. These three constraints must dominate the faithfulness constraint that opposes lengthening, \textsc{Dep-µ}:

\begin{tabular}{|c|c|c|c|}
\hline
/kana/ & \textsc{Ft-Bin} & \textsc{Non-Fin} & \textsc{Head (PrWd)} & \textsc{Dep-µ} \\
\hline
a. → (’ka:na) & & & 1 & \\
\hline
b. (’ka:na) & 1 W & & L & \\
\hline
c. (ka’na) & 1 W & & L & \\
\hline
d. kana & 1 W & & L & \\
\hline
\end{tabular}

The ranking in (9) explains why Foot-Binarity is satisfied by vowel lengthening in cases of the unstressable word syndrome: lengthening is the only way to achieve a binary foot while respecting Non-Finality, and \textsc{Headedness(PrWd)} demands that some foot be present.

Kager’s P-OT analysis also accounts for why iambic lengthening occurs. He assumes a constraint \textsc{Uneven-Iamb} that is violated by any foot except a light-heavy one. Iambic lengthening is the result of ranking \textsc{Uneven-Iamb} over \textsc{Dep-µ}, as in tableau (10). \textsc{Parse-Syllable}, which favors more complete footing, must also dominate \textsc{Dep-µ}, because simply positing fewer feet would be a way of avoiding the necessity of lengthening their head syllables.

\begin{tabular}{|c|c|c|c|c|c|}
\hline
/nemokotono/ & \textsc{Ft-Bin} & \textsc{Non-Fin} & \textsc{Head (PrWd)} & \textsc{Uneven-Iamb} & \textsc{Parse-Syll} & \textsc{Dep-µ} \\
\hline
a. → (ne’mo:)(ko’to:(no) & & & & 1 & 2 & \\
\hline
b. (ne’mo:)(ko’to:no & & & 2 W & 1 & L & \\
\hline
c. (ne’mo:)(ko’tono & & & 3 W & 1 & L & \\
\hline
\end{tabular}

Tableau (10) shows \textsc{Uneven-Iamb} ranked below \textsc{Non-Finality(foot)} and \textsc{Headedness(PrWd)} because (’ka:na) violates it but *(ka’na) and *kana do not.

Finally, this P-OT analysis accounts for why lengthening is not allowed in stranded penults, such as the third syllable of *(a’tʃo:)(’wo:wo. The crucial ranking puts \textsc{Uneven-Iamb}, which according to Kager (’wo: violates, over \textsc{Parse-Syllable}. This rules out the creation of a monosyllabic foot to achieve fuller foot parsing, as shown in tableau (11).

\begin{tabular}{|c|c|c|c|c|c|}
\hline
/afʃowowo & \textsc{Ft-Bin} & \textsc{Non-Fin} & \textsc{Head (PrWd)} & \textsc{Uneven-Iamb} & \textsc{Parse-Syll} & \textsc{Dep-µ} \\
\hline
a. → (a’tʃo:)(’wo:wo & & & & 2 & 1 & \\
\hline
b. (a’tʃo:)(’wo:wo & & & 1 W & 1 L & 2 W & \\
\hline
c. (a’tʃo:)(’wo:wo & 1 W & & 1 W & 1 L & 1 & \\
\hline
d. (a’tʃo:)(’wo:wo & 1 W & & L & 2 W & \\
\hline
\end{tabular}

In P-OT, the interaction between operations (in \textsc{Gen}) and constraints (in \textsc{Eval}) is
unidirectional: the operations apply, singly and in all combinations, and then the most harmonic result is chosen. From (9)–(11), it might seem that this is the only way to solve the problem of how operations and constraints interact in Hixkaryana and perhaps in CLIs generally. As we will now show, however, parallelism is not essential when constraints are violable.

4 Cross-level interactions in Harmonic Serialism

Because it inherits OT’s violable constraints, HS offers a successful and fully formalized solution to the problem of how constraints and operations interact in a derivation — a problem that could not be solved in earlier serial theories that combined rules with inviolable constraints. (For the background necessary to follow this argument, see McCarthy (this volume).)

We will illustrate this claim with an HS analysis of Hixkaryana. GEN includes two operations that are relevant to this analysis: footing and vowel lengthening. The vowel lengthening operation was defined in (5). Following Pruitt (2010), the footing operation is defined so as to parse any pair of adjacent unfooted syllables as a headed binary foot or any single unfooted syllable as a headed unary foot. The full candidate set includes the unchanged faithful candidate and the various results of applying any one of the operations once. For simplicity, we will ignore candidates with obviously fatal constraint violations, such as candidates with disyllabic trochaic feet.

We will also replace Kager’s ad hoc constraint UNEVEN-IAMB with the standard constraint SWP, which was mentioned in section 2. SWP is violated by any stressed light syllable, and thus it accounts for the widespread process of lengthening under stress, including iambic lengthening. In contrast, UNEVEN-IAMB disfavors even (ˈCVː) iambs relative to (CVˈCVː) ones, wrongly predicting the existence of languages that epenthesize a vowel to change (ˈCVː) into (VˈCVː).

Because the mapping /kana/ → (ˈka):na involves two distinct operations in GEN, stress assignment and lengthening, it has to pass through an intermediate step. That intermediate step is (ˈka)na, with a degenerate foot. For (ˈka)na to win at step 1, NON-FINALITY and HEEDEDNESS(PrWd) have to dominate FOOT-BINARITY and SWP, as shown in (12). (The reason for ranking FOOT-BINARITY over SWP will emerge in (16) and (18).)

(12) Step 1 of /kana/ → (ˈka):na in HS

<table>
<thead>
<tr>
<th>/kana/</th>
<th>NON-FIN (PrWd)</th>
<th>HEAD (PrWd)</th>
<th>FT-BIN</th>
<th>SWP</th>
<th>DEP-µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (ˈka)na</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. kana</td>
<td></td>
<td>1 W</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c. (kaˈna)</td>
<td>1 W</td>
<td>L</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. kana</td>
<td>1 W</td>
<td>L</td>
<td>L</td>
<td>1 W</td>
<td></td>
</tr>
</tbody>
</table>

The ranking in (12) is similar but not identical to the ranking required for the P-OT analysis in (9). The main difference is this: FOOT-BINARITY is surface-unviolated in Hixkaryana, so it is undominated in P-OT, which knows no other output level of representation than surface structure. In HS, however, FOOT-BINARITY must be dominated because it is is crucially violated at the first step of the /kana/ → (ˈka):na → (ˈka):na derivation. This is therefore an instance of what we referred to in the introduction as violation of the surface-true.

At step 2, the candidate (ˈka):na becomes available, because it is one operation away from
the step 1 output/step 2 input (/ˈkaː)na/. It satisfies Foot-Binarity and SWP at the expense of violating lower-ranking Dep-μ:

(13) Step 2 of /kana/ → (/ˈkaː)na in HS

<table>
<thead>
<tr>
<th>('ka)na</th>
<th>NON-FIN</th>
<th>HEAD (PrWd)</th>
<th>FT-BIN</th>
<th>SWP</th>
<th>Dep-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → ('ka):na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b. ('ka):na</td>
<td></td>
<td>1 W</td>
<td>1 W</td>
<td></td>
<td>L</td>
</tr>
</tbody>
</table>

This derivation converges at step 3 because no footing or lengthening operation will produce any further improvement:

(14) Step 3 of /kana/ → (/ˈkaː)na in HS — Convergence

<table>
<thead>
<tr>
<th>('kaː)na</th>
<th>NON-FIN</th>
<th>HEAD (PrWd)</th>
<th>FT-BIN</th>
<th>SWP</th>
<th>Dep-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → ('kaː)na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ('kaː)(na)</td>
<td></td>
<td>1 W</td>
<td>1 W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ('kaː)na:</td>
<td></td>
<td></td>
<td></td>
<td>1 W</td>
<td></td>
</tr>
</tbody>
</table>

The derivation of /nemokotono/ also involves footing and lengthening steps. At step 1, the first two syllables are parsed into a foot, satisfying Headedness(PrWd). (Footing proceeds from left-to-right because of the constraint Align-L(ft, wd) (McCarthy and Prince 1993a), not shown here.)

(15) Step 1 of /nemokotono/ → (neˈmoː)(koˈtoː)no in HS

<table>
<thead>
<tr>
<th>/nemokotono/</th>
<th>NON-FIN</th>
<th>HEAD (PrWd)</th>
<th>FT-BIN</th>
<th>PARSE-SYLL</th>
<th>SWP</th>
<th>Dep-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (neˈmo)kotono</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. nemokotono</td>
<td></td>
<td>1 W</td>
<td>5 W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This is another instance of violation of the surface-true because the winning candidate violates SWP, a constraint that is never violated by surface forms of this language.

At step 2 of this derivation, a second foot is built next to the first one, further improving performance on Parse-Syllable. Adding a second foot is preferred to lengthening the stressed vowel in the foot already built because Parse-Syllable dominates SWP:

(16) Step 2 of /nemokotono/ → (neˈmoː)(koˈtoː)no in HS

<table>
<thead>
<tr>
<th>(neˈmo)kotono</th>
<th>NON-FIN</th>
<th>HEAD (PrWd)</th>
<th>FT-BIN</th>
<th>PARSE-SYLL</th>
<th>SWP</th>
<th>Dep-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (neˈmo)(koˈto)no</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>b. (neˈmo)kotono</td>
<td></td>
<td>3 W</td>
<td>1 L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. (neˈmo):(koˈto)no</td>
<td></td>
<td>3 W</td>
<td>L</td>
<td>1 W</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Once again, the surface-unviolated constraint SWP is violated at an intermediate derivational step.

At step 3, one of the stressed vowels lengthens. The derivation produces the same result
regardless of which one lengthens first. (This is therefore a *convergent tie* — see section 4 of McCarthy (this volume).) In tableau (17), we take the arbitrary option of lengthening the stressed vowel on the left first:

\[(17) \quad \text{Step 3 of } /\text{nemokotono}/ \rightarrow (\text{ne}ˈ\text{mo}):(/\text{ko}ˈ\text{to}:):\text{no} \text{ in HS} \]

<table>
<thead>
<tr>
<th></th>
<th>NON-FIN</th>
<th>HEAD</th>
<th>FT-BIN</th>
<th>PARSE-SYLL</th>
<th>SWP</th>
<th>DEP-µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( \rightarrow (\text{ne}ˈ\text{mo}):(/\text{ko}ˈ\text{to}:):\text{no} )</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b. ( (\text{ne}ˈ\text{mo}):(/\text{ko}ˈ\text{to}:):\text{no} )</td>
<td></td>
<td>1</td>
<td>2 W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The other stressed vowel lengthens at step 4, and then the derivation converges at step 5.

The derivation of /atʃowowo/ proceeds as follows: /atʃowowo/ \( \rightarrow (aˈʧo)\text{owo} \rightarrow (aˈʧoː)\text{owo} \). Because the step 1 tableau involves the same constraint interactions as (15), we will go straight to step 2, which is where the most interesting candidates are found:

\[(18) \quad \text{Step 2 of } /\text{atʃowowo}/ \rightarrow (aˈʧo)\text{owo} \text{ in HS} \]

<table>
<thead>
<tr>
<th></th>
<th>NON-FIN</th>
<th>HEAD</th>
<th>FT-BIN</th>
<th>PARSE-SYLL</th>
<th>SWP</th>
<th>DEP-µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( \rightarrow (aˈʧo)\text{owo} )</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b. ( (aˈʧo)\text{owo} )</td>
<td>1 W</td>
<td></td>
<td>1 L</td>
<td>2 W</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td>c. ( (aˈʧo)(ˈwo)\text{wo} )</td>
<td>1 W</td>
<td></td>
<td>L</td>
<td>2 W</td>
<td></td>
<td>L</td>
</tr>
</tbody>
</table>

At step 2, FOOT-BINARITY blocks the creation of the degenerate foot (ˈwo). Because further foot parsing is foreclosed by FOOT-BINARITY and NON-FINALITY, iambic lengthening occurs instead, in response to the ranking of SWP above DEP-µ. Finally, the derivation converges at step 3.

As promised, this analysis illustrates the HS solution to the problem of how an active constraint can come to be violated in a derivation. Instead of relaxing serialism and allowing operations to apply simultaneously, as in P-OT or the repair-strategy analysis of Mester (1994), HS relies on OT’s ranked and violable constraints. In OT, from the fact that FOOT-BINARITY is active in forcing lengthening in /ˈkaːna/, it does not follow that it is inviolable. The potential for violability allows for the creation of a degenerate foot if that is required by a higher ranked constraint, like NON-FINALITY and HEADEDNESS(PrWd) in this analysis. The challenging conspiratorial interaction between blocking and repair is also dealt with straightforwardly in this serial theory. The constraint ranking produces the effect of the conspiracy by choosing which operation provides the best way of satisfying FOOT-BINARITY in a given context: by not footing at all when HEADEDNESS(PrWd) is already satisfied, as in (aˈʧo)\text{owo}, and by lengthening when NON-FINALITY and HEADEDNESS(PrWd) have forced the creation of a degenerate foot, as in /ˈkaːna/.

This analysis also explains why the lengthening repair is available for violations of SWP and HEADEDNESS(PrWd), but not in *(aˈʧo):(ˈwo)\text{wo}*. Because PARSE-SYLLABLE dominates SWP, (CV\text{'CV}) iambs are created in the course of the derivation, but they later change to (CV\text{'CV}) because SWP itself dominates DEP-µ. And because HEADEDNESS(PrWd) dominates FOOT-BINARITY, CVCV words are parsed with a degenerate foot (ˈCV)CV, which later becomes (CV)CV because FOOT-BINARITY itself dominates DEP-µ. There is no lengthening in *(aˈʧo):(ˈwo)\text{wo} because there is nothing to repair: (aˈʧo):(ˈwo)\text{wo} is not optimal earlier in the derivation, because PARSE-SYLLABLE is ranked below FOOT-BINARITY. Except when higher-ranking HEADEDNESS(PrWd) forces them, violations of FOOT-BINARITY never arise in the
course of the derivation. In HS, temporary violation of the surface-unviolated constraints SWP and FOOT-BINARITY is controlled by the grammar in the same way that P-OT controls violation of surface-violated constraints. Significantly, the HS analysis accounts for these facts without the dubious constraint UNEVEN-IAMB (violated by (ˈwo) as well as (neˈmo) and (koˈto)) that is required in the P-OT analysis.

Finally, the Hixkaryana analysis illustrates another of the core properties of HS highlighted in the introduction: full availability of structural operations. In tableaux (16) and (18), for example, lengthening a stressed vowel competes against building another foot. Because PARSE-SYLLABLE dominates SWP, foot-building is favored until no more feet can be built. Operations at different levels of the hierarchy are always available in GEN, and their applicability is controlled as usual by the constraint hierarchy. Thus, structure-building in HS is not strictly bottom-up (or top-down, for that matter). This has relevance to Prince & Smolensky’s argument against bottom-up constructionism in Tongan, for which see Elsman (this volume).³

Cross-level interactions are a special case of the general problem of formalizing the interaction between constraints and operations. There is nothing in the theory that would lead us to expect that footing could not create a violation of a constraint, or that lengthening could not apply afterwards to resolve it. By solving the problem of formalizing temporary violation and repair, HS also solves this problem of constraint/operation interaction, and maybe all others. Parallelism is not required.

5 Violation of the surface-true

The analysis of Hixkaryana in section 4 relies on an inherent property of HS: derivations may involve temporary violation of surface-unviolated constraints. Two other putative arguments for parallelism in the literature can also be reanalyzed by relying on this aspect of HS: the interaction of stress and syllable weight in Latin (Mester 1994: 15); and the interaction of syllabification and segmental processes in Lardil (Prince and Smolensky 1993/2004: 148-149).

5.1. Latin stress and syllable weight

Like Hixkaryana, Latin has CLIs between vowel length and metrical structure. The processes involved are known as iambic and cretic shortening, because the underlying forms of the words they affect look like the iambic (LH) and cretic (HLH) feet of Greek meter. Both processes are primarily associated with pre-Classical, Plautine Latin and are usually assumed to reflect a more colloquial speech style than Classical Latin.

In iambic shortening, a final long vowel in a LH disyllable is shortened:

(19) Iambic shortening⁹

<table>
<thead>
<tr>
<th>Underlying</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>amo:</td>
<td>'a.mo</td>
</tr>
<tr>
<td>puta:</td>
<td>'pu.ta</td>
</tr>
<tr>
<td>wolo:</td>
<td>'wo.lo</td>
</tr>
<tr>
<td>wiri:</td>
<td>'wi.ri</td>
</tr>
<tr>
<td>homo:</td>
<td>'ho.mo</td>
</tr>
</tbody>
</table>

HH disyllables, such as ˈlau.do: ‘I praise’, do not undergo shortening.

In cretic shortening, a final long vowel in a HLH trisyllable is shortened:
Cretic shortening

Underlying | Surface
---|---
diːkiːtoː | 'diː.ki.to
impeːra | 'im.pe.ra
mentioː | 'men.ti.o
desinoː | 'deː.si.no

LLH trisyllables, such as 'stu.de.oː ‘I study’, do not undergo shortening.

Mester (1994) argues that both of these shortening processes improve metrical parsing. He argues that the Latin foot is a maximally and minimally bimoraic trochee, (ˈLL) or (ˈH). Shortening in /putaː/ → (ˈpu.ta) yields a foot that is neither too small like *ˈpu.taː nor too large like *ˈpu.ta. Cretic shortening is a response to this same imperative, under the assumption that (ˈdiː)(ˌki.to) has a (ˌLL) foot after the primary stress.

Prince & Smolensky (1993/2004: section 4.5) show how Mester’s ideas can be expressed in P-OT. They begin with the basic stress pattern, which puts stress on a heavy penult, else the antepenult, in words of sufficient length:

<table>
<thead>
<tr>
<th>Weight Pattern</th>
<th>Main-stress foot</th>
<th>Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>...ˈLLσ</td>
<td>susˈ(pi.ki)a</td>
<td>‘suspicions (nom./acc.)’</td>
</tr>
<tr>
<td>...ˈHLσ</td>
<td>deː.klarˈ(ra)ti.o</td>
<td>‘declaration (nom.)’</td>
</tr>
<tr>
<td>...LˈHσ</td>
<td>aˈmiːke</td>
<td>‘friend (voc.)’</td>
</tr>
<tr>
<td>...HˈHσ</td>
<td>serˈ(mo)ne</td>
<td>‘speech (abl.)’</td>
</tr>
</tbody>
</table>

The foot is obviously trochaic. It is attracted to the right edge of the word, but subject to the requirement that the final syllable not be footed. This means that NON-FINALITY(ˈft), which is violated if the main-stress foot is final in the prosodic word, dominates ALIGN-R(ˈσ, word), which is fully satisfied only by final stress:

<table>
<thead>
<tr>
<th>/suspikia/</th>
<th>NON-FINALITY(ˈft)</th>
<th>ALIGN-R(ˈσ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → susˈ(pi.ki)a</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>b. sus.piˈ(ki.a)</td>
<td>1 W</td>
<td>1 L</td>
</tr>
</tbody>
</table>

Iambic shortening shows that NON-FINALITY(ˈft) is itself crucially dominated. Underlying /amoː/ becomes (ˈa.mo) because there is no other allowable way of parsing this word into a well-formed trochaic foot. The candidate (ˈa.mo) in (24b) fatally violates the Weight-to-Stress Principle of Prince (1990), which we formulate here as a constraint on foot structure, mimicking the notion of quantity-sensitivity in Hayes (1980):

WSP-FOOT

Assign a violation mark for any heavy syllable in the weak branch of a foot.

The candidate (ˈa)mo in (24c) is ruled out by FOOT-BINARITY, because the (ˈa) foot is monomoraic. The next candidate, aˈmo in (24d), has main stress on the final syllable in contravention of a different NON-FINALITY constraint, NON-FINALITY(ˈσ). Finally, footless (24e) is out because of the requirements that every lexical word correspond to a prosodic word (LX反映Pr) and that every prosodic word contain a head foot (HEADEDNESS(PrWd) (Selkirk 1995)). Ranking these markedness constraints above NON-FINALITY(ˈft) and the faithfulness constraint MAX-µ produces the desired result:
Iambic shortening in P-OT (after Prince and Smolensky 1993/2004: section 4.5)

<table>
<thead>
<tr>
<th>/amo/</th>
<th>WSP-Foot</th>
<th>FT-Bin</th>
<th>NON-FIN('σ)</th>
<th>LX ≈ PrHD(Pr)</th>
<th>NON-FIN('ft)</th>
<th>MAX-µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (‘a.mo)</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (‘a.mo):</td>
<td>1 W</td>
<td></td>
<td></td>
<td>1</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c. (‘a)mo:</td>
<td>1 W</td>
<td></td>
<td>L</td>
<td></td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>d. a(mo:)</td>
<td>1 W</td>
<td></td>
<td>1</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. a.mo:</td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

Prince & Smolensky’s analysis of cretic shortening is similar, except that shortening affects a foot that is assumed to follow the main stress in cretic words: /diːkitoː/ → (ˈdiː)(ˌki.to). The presence of this foot is an effect of PARSE-SYLLABLE, which rules out (ˈdiː)ki.to: and (ˈdiː)ki(to) in (25c,d). As in the case of iambic shortening, failure to shorten in (ˌki.to) violates WSP-Foot (see (25b)):

Cretic shortening in P-OT (after Prince and Smolensky 1993/2004: section 4.5)

<table>
<thead>
<tr>
<th>/dikito:/</th>
<th>WSP-Foot</th>
<th>PARSE-SYLL</th>
<th>MAX-µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (ˈdii)(ˌki.to)</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. (ˈdii)(ˌki.to)</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c. (ˈdii)ki.to;</td>
<td>2 W</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>d. (ˈdii)ki(to);</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

To complete this analysis of Cretic shortening, Prince & Smolensky require one additional constraint and ranking. A constraint against (‘HL) trochees, dubbed RhHRM (for rhythmic harmony) by Prince & Smolensky (1993/2004: 70-71), also dominates MAX-µ:

Cretic shortening in P-OT — RhHRM ≫ MAX-µ

<table>
<thead>
<tr>
<th>/dikito:/</th>
<th>RhHRM</th>
<th>MAX-µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (ˈdii)(ˌki.to)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b. (ˈdii)ki(to);</td>
<td>1 W</td>
<td>L</td>
</tr>
</tbody>
</table>

This will suffice for now as a summary of the P-OT analysis. Turning now to HS, we will apply the same analytic strategy seen in the analysis of Hixkaryana: foot parsing creates a foot with marked quantitative structure that is made less marked by changing quantity at a later stage of the derivation. In the case of Latin, the (‘LH) trochee that is initially created violates WSP-Foot, which is satisfied later in the derivation when MAX-µ is violated.

We will enter the derivation at what might be called step 2, after syllabification. Tableau (27) shows step 2 of the derivation of /amoo/. GEN can build any single foot or shorten a vowel, but not both. Failure to build a foot, as in (27d), is ruled out by LX≈Pr and HEADEDNESS(PrWd). Building a foot on just the final or initial syllable, as in (27b,c), violates FOOT-BINARITY or NON-FINALITY(‘σ). The crucial move in (27) is ranking WSP-Foot below FOOT-BINARITY and NON-FINALITY(‘σ). This ranking allows (ˈa.mo:) to win at step 2, setting the stage for shortening at step 3.
Step 2 from /amoː/: NONFIN(ˈσ), FOOT-BINARITY ≫ WSP-Foot

<table>
<thead>
<tr>
<th></th>
<th>NON-FIN(ˈσ)</th>
<th>Ft-Bin</th>
<th>LX ≈ PR-Hd(Pr)</th>
<th>NON-FIN(ˈft)</th>
<th>WSP-Foot</th>
<th>ALIGN-R(ˈσ)</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>('a.moː)</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>a('moː)</td>
<td>1 W</td>
<td></td>
<td>1</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>('a)moː</td>
<td>1 W</td>
<td></td>
<td>L</td>
<td>L</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>a.moː</td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

This is a case of violation of the surface-true — Latin has no surface (ˈLH) feet, because WSP-Foot is surface-unviolated.

Shortening occurs at step 3 because WSP-Foot dominates MAX-μ, as in the P-OT analysis:

Step 3 from /amoː/: WSP-Foot ≫ MAX-μ

<table>
<thead>
<tr>
<th></th>
<th>NON-FIN(ˈft)</th>
<th>WSP-Foot</th>
<th>ALIGN-R(ˈσ)</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>('amo)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>('amo)</td>
<td>1 1 W</td>
<td>1</td>
<td>L</td>
</tr>
</tbody>
</table>

Shortening occurs at this step because all of the high-ranking foot-parsing constraints have been satisfied. This allows EVAL to turn its attention to the quantitative constraint WSP-Foot. The result of shortening is that ('amo) now conforms with expectations about Latin surface forms, and indeed the derivation converges at the next step.

The HS derivation of cretic shortening includes two steps of foot construction before shortening occurs: /desinə/ → deː.si.noː → (ˈdeː)si.noː → (ˈdeː)(si.noː) → (ˈdeː)(ˌsi.no). At the first post-syllabification step, step 2, the heavy antepenult is parsed into a monosyllabic foot that receives main stress. This candidate wins because the candidates with more complete footing, (29b,c), violate RHRM or NON-FINALITY(ˈft), both of which dominate PARSE-SYLLABLE:

Step 2 from /desinə/: — NON-FINALITY(ˈft) ≫ PARSE-SYLL

<table>
<thead>
<tr>
<th></th>
<th>NON-FIN(ˈσ)</th>
<th>RHRM</th>
<th>NON-FIN(ˈft)</th>
<th>WSP-Foot</th>
<th>ALIGN-R(ˈσ)</th>
<th>PARSE-SYLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>('deː)si.noː</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>b.</td>
<td>('deː)si)noː</td>
<td>1 W</td>
<td></td>
<td></td>
<td>2</td>
<td>1 L</td>
</tr>
<tr>
<td>c.</td>
<td>deː('si.no)</td>
<td>1 W</td>
<td>1 W</td>
<td>1 L</td>
<td>1 L</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>deː.si(ˈno)</td>
<td>1 W</td>
<td>1 W</td>
<td>L</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Candidate (29c) involves a ranking disjunction — NON-FINALITY(ˈft) or WSP-Foot dominates ALIGN-R(ˈσ, word) and PARSE-SYLLABLE — but this disjunction will be partly resolved in a moment, when we see that PARSE-SYLLABLE has to dominate WSP-Foot.

At step 3, the options for further parsing of (ˈdeː)si.no: include all of the ways of adding a foot to the last two syllables, plus shortening one of the vowels or doing nothing. Shortening at this point would gratuitously violate MAX-μ, while doing nothing violates PARSE-SYLLABLE. To get the desired parse (ˈdeː)(ˌsi.no) at this step, PARSE-SYLLABLE has to dominate WSP-Foot:
Step 3 from /desinoː/ — PARSE-SYLLABLE ≫ WSP-FOOT

<table>
<thead>
<tr>
<th></th>
<th>PARSE-</th>
<th>WSP-</th>
</tr>
</thead>
<tbody>
<tr>
<td>('de);si:no;</td>
<td>SYLL</td>
<td>FOOT</td>
</tr>
<tr>
<td>a. →</td>
<td>('de);(si:)</td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>('de);si:</td>
<td>2 W L</td>
</tr>
<tr>
<td>c.</td>
<td>('de);si(</td>
<td>no;</td>
</tr>
</tbody>
</table>

As in the derivation of ('amo), the surface unviolated constraint WSP-FOOT is violated at this intermediate step.

Next, at step 4, input ('de);(si:no;) undergoes shortening of the final vowel because WSP-FOOT dominates MAX-μ:

Step 4 from /desinoː/

<table>
<thead>
<tr>
<th></th>
<th>WSP-</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>('de);(si:no;)</td>
<td>FOOT</td>
<td></td>
</tr>
<tr>
<td>a. →</td>
<td>('de);(si:no;)</td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>('de);(si:no;)</td>
<td>1 W L</td>
</tr>
</tbody>
</table>

The derivation converges at step 5.

WSP-FOOT is surface-unviolated in this variety of Latin, and it is therefore undominated in the P-OT analysis. In the HS analysis, however, it is violated at an intermediate step of the iambic and cretic shortening derivations. This is reflected in a difference in ranking: WSP-FOOT has to be dominated by NON-FINALITY('σ'), FOOT-BINARITY, and PARSE-SYLLABLE in the HS analysis, but these four constraints are surface-unviolated and therefore unrankable in P-OT. Thus, although ('LH) trochees are absent from surface forms, they emerge and are later eliminated by shortening in the course of the derivation. This is therefore an instance of violation of the surface-true.

In terms of empirical coverage, the analysis in HS presented here and the analysis in P-OT in Prince & Smolensky (1993/2004) are essentially equivalent, but there is an interesting difference between the constraints required for each analysis. With the constraints discussed thus far, the P-OT analysis cannot account for why /desinoː/ shortens the final vowel, yielding ('de);(si:no;), instead of shortening the antepenultimate vowel, yielding *(('de.si)(no;). Prince & Smolensky resolve this tie by calling on the constraint PkPROM, which favors having a heavy main-stressed syllable. More recent developments in OT offer a couple of other options: some sort of contiguity constraint favoring peripheral over medial shortening (Kienstowicz 1994b, McCarthy and Prince 1995, 1999); or a positional faithfulness constraint discouraging shortening in the main-stressed syllable (Beckman 1998).

This issue does not arise at all in the HS analysis, however. As is apparent from tableaux (29)–(31), *(('de.si)(no;) is never even a candidate, much less a serious challenger to the intended winner. It is not a candidate because it would require an intermediate step with a ('HL) trochee, ('de.si). Undominated RHHRM ensures that such feet never win even at intermediate derivational steps. The HS analysis therefore has an advantage over the P-OT analysis: it can explain, and need not stipulate, which syllable is affected by cretic shortening.

5.2. Lardil syllabification and augmentation

architecture for OT from the interaction of syllabification and augmentation in Lardil.

Lardil limits its coda consonants to coronal sonorants, plus labial and dorsal nasals that share place with the following onset. Lardil has a process of augmentation that affects unsuffixed CVC roots, making them disyllabic for word-minimality reasons. If the final consonant of the CVC root is a sonorant, augmentation adds C, where C is a stop that is homorganic with the preceding consonant (32a). Roots ending in obstruents augment by adding only a, with the root-final consonant syllabified as an onset (32b). The a augment is also used with r-final roots (32c), because the language prohibits clusters of r followed by a coronal (except lamino-alveolars) (Hale and Nash 1997: 250).

(32) Lardil augmentation (all words nominative case, data from Hale et al. (1981))

a. CV augmentation after sonorants (except r)
   /wun/ wun.ta ‘rain’
   /maŋ/ maŋ.ta ‘hand’
   /ŋ/ ŋ.ʈa ‘neck’
   /kaŋ/ kaŋ.k ‘speech’

b. V augmentation after obstruents
   /pat/ pa.ta ‘west’
   /jak/ ja.ka, *jak.k ‘fish’

c. V augmentation after r
   /teɾ/ teɾ.a, *teɾ.ta ‘thigh’

Why do Prince and Smolensky see these data as challenging for serialism? They consider a serial theory in which “Syllabification” and “Augmentation” are rules whose order must be fixed, and they argue that Syllabification must precede Augmentation to distinguish the forms in which V is added from those that get CV: “Augmentation inserts an onset only when the stem-final consonant is already parsed as a coda” (p. 148). Thus, the serial derivation as they conceive it must go something like this:


```
/wun/ Syllabification [wun]o /jak/
```

Prince and Smolensky’s argument against serialism relies on the observation that g codas must be followed by k, so g’s [dorsal] place is licensed by association with the following onset: “Now consider /kan/. When Syllabification applies, the g is no more syllabifiable as a coda than the k of /yak/ [our /jak/] — g can only be a coda when linked to a following onset. The situation is obviously not improved by trying to allow Augmentation to precede Syllabification… In short, Syllabification and Augmentation are mutually interdependent: each ‘triggers’ the other.”

Although this is a valid argument against the serial theory implicit in (33), HS is not such a theory. The theory in (33) has inviolable constraints — because free-standing g is never a licit coda in surface forms, it cannot be parsed as a coda in intermediate kan — and it presupposes that epenthesis and (re-)syllabification can never occur in parallel. This is not HS. In HS, as in OT generally, constraints are violable, and even surface-unviolated constraints may be violated in the course of a derivation.

A further point of difference between HS, as we understand it here, and the serial theory implied by (33) is that research in HS, as in OT generally, takes the nature of GEN to be an empirical hypothesis. Although we may speak of HS’s GEN “performing only one operation at a time”, the question of what a single operation can do is an empirical one. Another way to put
this question is to ask exactly how much parallelism is needed, or how much must a single operation in GEN be allowed to do. An example of this sort of reasoning can be found in McCarthy’s (this volume) argument that resyllabification must occur in parallel with syllable-altering processes like epenthesis or syncope.

Both of these matters — violation of the surface-true and the question of GEN’s power — are important in understanding Lardil augmentation in HS. A typical HS derivation for Lardil will look like this:

(34) HS derivation of wun.ta

/\wun/ \rightarrow wun \rightarrow ('wun) \rightarrow ('wu.na) \rightarrow ('wun.Ca) \rightarrow ('wun.ta)

The symbol C stands for a placeless plosive. At step 1 of this derivation, /\wun/ is syllabified as the single syllable wun, and at step 2 (shown in tableau (35)) it is parsed into a foot to satisfy HEADEDNESS(PrWd), at the expense of violating FOOT-BINARITY. (Lardil codas are non-moraic.)

(35) Step 2 of /\wun/ \rightarrow 'wun.ta

<table>
<thead>
<tr>
<th>wun</th>
<th>HD(Pr)</th>
<th>FT-BIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>('wun')</td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>wun</td>
<td>1 W</td>
</tr>
</tbody>
</table>

Augmentation occurs at step 3 because FOOT-BINARITY dominates DEP, as shown in tableau (36) (though see Moore-Cantwell (this volume) for a different view of such word-minimality effects). This tableau also shows that FOOT-BINARITY dominates the constraint ALIGN-R(MWORD, α), which is the trigger of Ca augmentation in Prince & Smolensky’s analysis. This constraint is violated by intermediate ('wu.na) because the rightmost consonant in the morphological word /\wun/ is not also rightmost in some syllable.

(36) Step 3 of /\wun/ \rightarrow 'wun.ta

<table>
<thead>
<tr>
<th>('wun')</th>
<th>FT-BIN</th>
<th>ALIGN-R</th>
<th>DEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>('wu.na')</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>('wun')</td>
<td>1 W</td>
<td>L</td>
</tr>
</tbody>
</table>

Step 4 of this derivation brings it into conformity with ALIGN-R(MWORD, α) by epenthesisizing a placeless plosive consonant C into onset position, pushing the n back into the coda where it belongs. The epenthized consonant is placeless because of the undominated faithfulness constraint DEP(place), which is violated if any more fully specified consonant is epenthized, as in candidates (37b,c,d). It is a plosive by emergence of the unmarked: plosives minimize sonority in the onset, as required by the margin hierarchy (Prince and Smolensky 1993/2004: 151ff.).
At the last pre-convergence step of the derivation, the [coronal] feature of \( n \) spreads into the following placeless C. This occurs because a markedness constraint favoring consonants with place (call it HAVE-PLACE) dominates any faithfulness constraint that militates against spreading place:

\[
(38)\quad \text{Step 5 of } /wun/ \rightarrow \acute{wun}.ta
\]

<table>
<thead>
<tr>
<th>( (\acute{\text{wun}}.Ca) )</th>
<th>HAVE-PLACE</th>
<th>“NO-SPREAD”</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( (\acute{\text{wun}} \text{ t}a) )</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. ( (\acute{\text{wun}} \text{ Ca}) )</td>
<td>1 W</td>
<td>L</td>
</tr>
</tbody>
</table>

HAVE-PLACE must be ranked below ALIGN-R(MWORD, \( \sigma \)), else it would block epenthesis of the placeless plosive C at step 4.

Augmentation with \( a \) rather than \( Ca \) occurs when the root-final consonant is an obstruent. Obstruent codas are illicit because of the constraint in (39), which is based on typological observations by Zec (1995) and others about the preference for high sonority in codas:

\[
(39)\quad *\text{CODA}(\text{obst})
\]

Assign a violation mark for every \([-\text{sonorant}]\) consonant in coda position.

Independent evidence for \(*\text{CODA}(\text{obst})\) in Lardil comes from the observation that a word-final coronal obstruent \(/t/\) or \(/\tilde{t}/\) changes into the corresponding rhotic \(/\tilde{r}/\) or \(/\tilde{\eta}/\) (Hale 1973: 426 fn. 32): /ja\tilde{r}put/ \( \rightarrow \) ja\tilde{r}pu\tilde{r} ‘snake, bird’.

Now, suppose the derivation of \(/\text{pat}/ \rightarrow \acute{\text{pa}}.ta\) has proceeded through step 3, yielding \( (\acute{\text{pa}}.ta) \). At step 4, shown in tableau (40), \( (\acute{\text{pat}}.Ca) \) is a candidate, but it loses because \(*\text{CODA}(\text{obst})\) dominates ALIGN-R(MWORD, \( \sigma \)). This is therefore the convergence step for the \(/\text{pat}/ \rightarrow \acute{\text{pa}}.ta\) derivation.
Now we are ready to tackle /kaŋ/ → 'kaŋ.ka, which is the basis of Prince & Smolensky’s argument for parallelism. In Lardil surface forms, coda ŋ must always be followed by onset k — although Lardil has nonhomorganic clusters like nk, it does not allow ŋt, and there are alternations showing that it deletes word-final ŋ. /wuŋkunuŋ/ → wuŋkunu ‘queen-fish’ (cf. wuŋkunuŋin ‘id. non-future accusative’). Except for [coronal], consonantal place features have to be licensed by association with onset position:

(41)  *CODA(lab/dors)

Assign a violation mark for every instance of the features [labial] or [dorsal] not associated with a consonant in onset position.

A place feature that is licensed in this way can also be associated with a preceding coda without violating this constraint (Goldsmith 1990: 123-128, Ito 1989), so coda ŋ is allowed only if followed by onset k. They share a single [dorsal] place feature that is associated with the onset and thereby avoids violating *CODA(lab/dors).

Although *CODA(lab/dors) is surface-unviolated, violation in the course of the derivation is also possible in HS, and that is how /kaŋ/ → 'kaŋ.ka is derived. Suppose that *CODA(lab/dors) is crucially dominated by PARSE-SEGMENT, which is violated once by every unsyllabified segment. It follows that /kaŋ/ will be syllabified as the single syllable kag at step 1, as shown in tableau (42). This tableau pits fully-syllabified (42a) against partially syllabified (42b), which loses because of its violation of top-ranked PARSE-SEGMENT. Candidate (42c) takes a different path entirely, deleting a segment rather than parsing a syllable. This too is a non-starter because of PARSE-SEGMENT.

(42)  Step 1 of /kaŋ/ → 'kaŋ.ka

<table>
<thead>
<tr>
<th>/kaŋ/</th>
<th>PARSE-SEG</th>
<th>HEAD</th>
<th>FT-BIN</th>
<th>*CODA(lab/dors)</th>
<th>DEP</th>
<th>MAX-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>(σ (\downarrow k a \tilde{\eta}))</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>(σ (\downarrow k a \tilde{\eta}))</td>
<td>1 W</td>
<td>1</td>
<td>L</td>
<td>1 W</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>ka</td>
<td>2 W</td>
<td>1</td>
<td>L</td>
<td>1 W</td>
<td></td>
</tr>
</tbody>
</table>

In tableau (42), all candidates violate HEADEDNESS(PrWd). This is unavoidable at step 1 if it is assumed that GEN cannot build syllables and feet at the same time, and feet cannot be built before syllables (see note 8). At step 2, satisfying this top-ranked constraint takes precedence over other markedness violations. Although there is also a violation of *CODA(lab/dors) to worry about, the ranking of HEADEDNESS(PrWd) over *CODA(lab/dors) gives priority to building a foot rather than repairing the coda ŋ. Tableau (43) certifies this argument formally.
From now on, we will return to the “flat” representation of syllable and foot structure used in earlier tableaux (35).

The input/faithful candidate at step 3 violates two active markedness constraints, FOOT-BINARITY and *CODA(lab/dors). Epenthesizing the augment vowel deals with both, making the foot binary and (temporarily) moving the ŋ out of coda position:

\[(44) \text{ Step 3 of } /\text{kaŋ}/ \rightarrow \text{ˈkaŋ.ka} \]

What happens next depends on the ranking of ALIGN-R(MWORD, ŋ). Previously, in tableau (36), we saw that it is crucially dominated by the surface-unviolated constraint FOOT-BINARITY. But it must itself dominate another surface-unviolated constraint, *CODA(lab/dors). This ranking forces (ˈka.ŋa) to map to (ˈkaŋ.Ca) at step 4:

\[(45) \text{ Step 4 of } /\text{kaŋ}/ \rightarrow \text{ˈkaŋ.ka} \]

Finally, at step 5, the [dorsal] feature of the ŋ spreads rightward to supply a place feature for the placeless plosive C. This is ensured by the ranking in (38) as well as by *CODA(lab/dors), which requires that the [dorsal] feature be licensed by association with an onset. Deletion of ŋ from the coda, as in (46c), is non-viable because C remains placeless.
Step 5 of /kaŋ/ → ᵇkagainst

<table>
<thead>
<tr>
<th>(ˈkaŋ.Ca)</th>
<th>*CODA (lab/dors)</th>
<th>HAVE-PLACE</th>
<th>MAX-C</th>
<th>“NO-Spread”</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (ˈkaŋ.ka)</td>
<td></td>
<td></td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>b. (ˈkaŋ.Ca)</td>
<td>1 W</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c. (ˈka.Ca)</td>
<td>1 W</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

This derivation then converges at step 6.

Due diligence requires us to show that this analysis of /kaŋ/ → ˈkaŋ.ka does not adversely affect ŋ deletion in polysyllables like /wuŋkunuŋ/ → wuŋkunu. In the derivation of ˈkaŋ.ka, epenthesis of a occurs not to license the ŋ but rather to satisfy Foot-Binarity; licensing of ŋ is a mere side effect. When Foot-Binarity is not at stake, unlicensed ŋ is deleted because DEP dominates MAX-C — a ranking that we have shown (but not used) throughout the analysis of Lardil. With this ranking, wuŋkunu beats *wuŋkununa and *wuŋkununaŋa never gets a chance to be a candidate.18

Lardil has another potential argument for parallelism: flap-final /teɾ/ augments as ˈte.ɾa, not *ˈte.ɾa. Hale attributes this to a surface-true prohibition against clusters of r+coronal. The question is how the constraint *r+coronal can rule out the candidate *te.ɾ.Ca at step 4 of the derivation.

Following Chomsky & Halle (1968: 336-337) and the literature on underspecification theory (e.g., Kiparsky 1982, Pulleyblank 1986), we might propose that a markedness constraint is violated by a form if and only if its structural description is nondistinct from that form. A structural description and a form are non-distinct if and only the form does not contradict any of the requirements in the structural description. The clusters rt and rC are nondistinct from each other and from the structural description of *r+coronal. But the clusters rp and ik are distinct from the structural description of *r+coronal, and so they are allowed. Considerations of this sort have not played much of a role in P-OT because P-OT’s markedness constraints evaluate only fully-specified surface representations. But HS allows for the possibility of underspecified intermediate representations, such as ter.Ca. It is therefore no surprise that HS should adopt a criterion for constraint applicability that recalls the pre-OT derivational literature.

Alternatively, GEN could be given the power to epenthesize a segment and spread place onto it in a single step. This approach conflates steps 4 and 5 in (45) and (46) into a single step, and it allows ˈte.ɾa and *ˈte.ɾ.ta to compete directly. Ideally, we would resolve the question of which approach is right, but we have exhausted the Lardil evidence bearing on this question. General techniques of OT and HS typology are applicable (Wolf and McCarthy 2009), but applying them now would divert us from our goals.

Returning to our main point, we have argued that Lardil exemplifies violation of the surface-true constraint *CODA (lab/dors) is violated in the course of the derivation of ˈkap.ka from /kaŋ/. Constraint ranking provides a principled account of when and how this violation arises and when and how it is eliminated, not only in /kaŋ/ → ˈkap.ka but also in /wuŋkunuŋ/ → wuŋkunu. Prince & Smolensky’s argument for parallelism is really an argument against a serial theory with inviolable constraints. It is a valid objection to rule+constraint theories like those discussed in section 2, but it is not a valid objection to HS.

Prince & Smolensky’s argument also shows the inadequacy of bottom-up serial theories. The serial derivation in (33) is one in which syllabification strictly precedes foot parsing, so the
augmentation step is absolutely respectful of the difference between codas (e.g., in \([\text{wun}]\)) and non-codas (e.g., in \([\text{ja}\,\sigma]\)) emerging from the syllabification step. As shown by augmentation-triggered resyllabification in tableau (36), HS is not strictly bottom-up. Indeed, that is the topic of the next section.

6 Harmonic Serialism is not bottom-up

Some arguments for parallelism in the literature rely on the assumption that serial theories are inherently bottom-up: the interaction of phonological phrasing and cliticization in English (McCarthy 2002: 146-149); the interaction of stress and sonorant syllabification in English (Pater 2000: 248-250); and the interaction of stress and syllabification in Tongan (Prince and Smolensky 1993/2004: 33-36). We discuss the two English cases in this section; for Tongan, see Elsman (this volume). As we will argue, HS has full availability of structural operations at every step of the derivation; thus, it is not bottom-up.

The prosodic structure assigned to a function word often depends on the larger context in which it finds itself. In English, for example, monosyllabic prepositions like to are reduced phrase-medially (47a) but unreduced phrase-finally (47b):

(47) Function word stress in English
   a. Unstressed phrase-medially
      John spoke to Bill. (pronounced to)
   b. Stressed phrase-finally
      Who did John speak to? (pronounced tú)
      Who did John speak to yesterday?

The general rule is that monosyllabic function words (other than object pronouns) are stressed and consequently unreduced at the end of a phonological phrase (PPh), but they are otherwise unstressed and reduced.

Following Selkirk (1995), we assume that an unstressed, reduced function word like to is not a prosodic word (PrWd) like \([\text{John}]_{\text{PWd}}\) or \([\text{spoke}]_{\text{PWd}}\). Instead, English function words are normally parsed as immediate constituents of the phonological phrase:

(48) Structure of reduced function word (after Selkirk 1995)
   \{[\text{John}]_{\text{PWd}}\}_{\text{PPh}} \{[\text{spoke}]_{\text{PWd}} \text{ to } [\text{Bill}]_{\text{PWd}}\}_{\text{PPh}}

As in this example, throughout this section we will use brackets to delimit prosodic words and braces for phonological phrases, omitting the subscripts. See Selkirk (1995) for evidence that this is the correct structure, rather than attachment of to to the preceding or following prosodic word.

Selkirk’s analysis relies on a combination of alignment constraints and constraints on the prosodic hierarchy. Among the alignment constraints are two that require a lexical word and a prosodic word to begin and end together:

(49) ALIGN-L(LexWd, PWd)
    Assign a violation mark for every lexical word whose left edge does not coincide with the left edge of a prosodic word.

(50) ALIGN-R(LexWd, PWd)
    Assign a violation mark for every lexical word whose right edge does not coincide with the right edge of a prosodic word.

These constraints are violated by structures where a function word is incorporated into a preceding or following prosodic word: \([\text{spoke }\text{ to }\text{ Bill}]\).

The constraints on the prosodic hierarchy include these two:
(51) PWdCON
Assign a violation mark for every prosodic word that does not contain a lexical word.

(52) EXHAUSTIVITY(PPh)
Assign a violation mark for every constituent of type lower than prosodic word that is immediately dominated by PPh.

PWdCON is violated by any stressed function word like \([tó]_{pwd}\). EXHAUSTIVITY(PPh) is a no-skipping constraint on the levels of the prosodic hierarchy: because the phonological phrase immediately dominates the prosodic word in the hierarchy, a structure is marked if it contains an instance of phonological phrase immediately dominating some lesser constituent, such as the foot or syllable. The syllable tô in (48) violates this constraint.

A P-OT grammar that will produce the desired output in (48) is given in (53). The candidates represent the four logically possible dispositions of to, according to Selkirk (1995): as a “free clitic” that is immediately dominated by phonological phrase (53a); as a free-standing prosodic word (53b); as a proclitic incorporated into the following prosodic word (53c); and as an enclitic incorporated into the preceding prosodic word (53d). Each of the four constraints in (49)–(52) is violated by one of these candidates. The winner is (53a) because it violates the lowest-ranking constraint, EXHAUSTIVITY(PPh).

(53) P-OT analysis of (John) spoke to Bill (after Selkirk 1995)

<table>
<thead>
<tr>
<th>spoke to Bill</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>PWdCON</th>
<th>EXH (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ([{\text{spoke} \ tô \ [\text{Bill}]}])</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b. ([{\text{spoke} \ [tô] \ [\text{Bill}]}])</td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
</tr>
<tr>
<td>c. ([{\text{spoke} \ [tô \ [\text{Bill}]}])</td>
<td>1 W</td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td>d. ([{\text{spoke} \ tô \ [\text{Bill}]}])</td>
<td>1 W</td>
<td></td>
<td></td>
<td>L</td>
</tr>
</tbody>
</table>

Thus far, nothing has been said about the distribution of phonological phrases. In Selkirk (1995), the boundaries of phonological phrases are determined by the following constraint:

(54) ALIGN-R(Lex^{max}, PPh)
Assign a violation mark for every maximal projection of a lexical category whose right edge does not coincide with the right edge of a phonological phrase.

This constraint is unviolated in English surface forms. It is responsible for the single phonological phrase in … \{spoke to Bill\}, and it is also responsible for the pair of phonological phrases in … \{speak to\} \{yesterday\}.

Phrase-internally, a function word is parsed as an immediate constituent of a phonological phrase, as shown in (53), but this option is ruled out when the function word is phonological phrase-final. The responsible constraint is one that requires every phonological phrase to end in a prosodic word:

(55) ALIGN-R(PPh, PWd)
Assign a violation mark for every phonological phrase whose right edge does not coincide with the right edge of a prosodic word.

With ALIGN-R(LexWd, PWd) also prohibiting a phonological phrase-final function word from encliticizing onto the preceding prosodic word, the only option left is to parse the function word as a prosodic word on its own. This is a violation of PWdCON, which must therefore be
The interaction illustrated in tableau (56) is a top-down effect. Whether a function word is parsed as a prosodic word depends on where it is situated in the phonological phrase, which is the next level of structure above the prosodic word. McCarthy (2002) uses this top-down effect to argue for P-OT over HS. The essence of the argument is this: phonological phrases are constructed out of prosodic words, so the prosodic words cannot be built after the phonological phrases. But (56) shows that prosodic words can’t be built without knowing where the boundaries of phonological phrases are. (That is why the crucial competitors in (56) include candidates that differ from one another in both prosodic word and phonological phrase structure.) There is no serial ordering of prosodic word construction and phonological phrase construction that can meet both of these requirements, so the prosodic words and phonological phrases need to be built simultaneously, by parallel rather than serial optimization.

This argument for P-OT goes through only under the assumption that the serial alternative is strictly bottom-up — specifically, it argues against a theory that builds prosodic word structure, then builds phonological phrase structure, but is unable to revise the prosodic word structure to reflect the changed conditions occasioned by the appearance of phonological phrase structure. This assumption is neither necessary nor even plausible in HS, however, as we now demonstrate.

We will first lay out the derivations and then explain how the grammar produces them. The derivation of (John) spoke to Bill proceeds as follows:

(57)  Derivation of (John) spoke to Bill
Step 1  [spoke] to Bill (or spoke to [Bill])
Step 2  [spoke] to [Bill]
Step 3  {[spoke] to [Bill]}
Step 4  {[spoke] to [Bill]} — Convergence

At steps 1 and 2, prosodic word structure is built. The grammar does not specify whether spoke or Bill is parsed into a prosodic word first, but by the end of step 2 both prosodic words have been created. (This is another example of a convergent tie — cf. tableau (17).) At step 3, every prosodic word that the grammar requires has been built, and the only remaining harmony-improving operation is parsing the VP into a phonological phrase. The output of step 4 is identical with the output of step 3, so the derivation terminates.

A grammar that will produce this derivation is given in (58)–(61). It includes all of the constraints and rankings that are in Selkirk’s P-OT analysis, plus two additional rankings that will be explained shortly. At step 1, the options include building a prosodic word (58a), doing nothing (58b), and building a phonological phrase (58c). Doing nothing is no option at all, because (58b) is harmonically bounded by the winner. Building a phonological phrase...
first (58c) is ruled out by ranking ALIGN-L/R(LexWd, PWd) above ALIGN-R(Lex\textsuperscript{max}, PPh). This is the first of the two rankings that are required in the HS analysis but are not required (though consistent with) the P-OT analysis:

(58)  Step 1 of HS analysis of *(John) spoke to Bill*

<table>
<thead>
<tr>
<th>spoke to Bill</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex\textsuperscript{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PW\textsubscript{D} CON</th>
<th>EXH (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → [spoke] to Bill</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. spoke to Bill</td>
<td>2 W</td>
<td>2 W</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. {spoke to Bill}</td>
<td>2 W</td>
<td>2 W</td>
<td>L</td>
<td>1 W</td>
<td>3 W</td>
<td></td>
</tr>
</tbody>
</table>

We are ignoring the convergent tie between (58a) and *spoke to [Bill]*.

At step 2, the other prosodic word is built in response to the same constraint interaction as step 1:

(59)  Step 2 of HS analysis of *(John) spoke to Bill*

<table>
<thead>
<tr>
<th>[spoke] to Bill</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex\textsuperscript{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PW\textsubscript{D} CON</th>
<th>EXH (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → [spoke] to [Bill]</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [spoke] to Bill</td>
<td>1 W</td>
<td>1 W</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. {[spoke] to Bill}</td>
<td>1 W</td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
<td>2 W</td>
<td></td>
</tr>
</tbody>
</table>

At step 3, the choice is between building a phonological phrase (60a), doing nothing (60b), or parsing the function word *to* as a prosodic word. The highest-ranking constraint that distinguishes among these candidates is ALIGN-R(Lex\textsuperscript{max}, PPh), and it favors building a phonological phrase.

(60)  Step 3 of HS analysis of *(John) spoke to Bill*

<table>
<thead>
<tr>
<th>[spoke] to [Bill]</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex\textsuperscript{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PW\textsubscript{D} CON</th>
<th>EXH (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → {[spoke] to [Bill]}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. [spoke] to [Bill]</td>
<td></td>
<td></td>
<td>1 W</td>
<td></td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c. [spoke] [to] [Bill]</td>
<td></td>
<td></td>
<td>1 W</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

Finally, at step 4, there are two main options, doing nothing (61a) or parsing *to* as a prosodic word (61b). As in Selkirk’s analysis, ranking PW\textsubscript{D} CON over EXHAUSTIVITY(PPh) disfavors the prosodic word parse of *to* in phrase-medial position, and so the derivation converges.
Step 4 of HS analysis of *John* spoke to Bill

<table>
<thead>
<tr>
<th></th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex_{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PWd CON</th>
<th>Exh (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>{{spoke} to [Bill]}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>{{spoke} [to] [Bill]}</td>
<td></td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
</tr>
</tbody>
</table>

In Selkirk's P-OT analysis, although function words are not normally parsed as prosodic words because PWdCON dominates EXHAUSTIVITY(PPh), a phonological phrase-final function word is parsed as a prosodic word because ALIGN-R(PPh, PWd) dominates PWdCON. Exactly the same interaction occurs in the HS analysis, but its effect is intrinsically ordered after construction of the phonological phrase. We take up the derivation of *(Who did John) speak to yesterday?* at step 3, when *speak* and *yesterday* have already been parsed as prosodic words. The winning candidate parses the prosodic word *yesterday* as a phonological phrase. It harmonically bounds all of its competitors, as tableau (62) shows.

Step 3 of HS analysis of *(Who did John) speak to yesterday?*

<table>
<thead>
<tr>
<th>[speak] to [yesterday]</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex_{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PWd CON</th>
<th>Exh (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>[speak] to {{yesterday}}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>[speak] to [yesterday]</td>
<td></td>
<td></td>
<td></td>
<td>2 W</td>
<td></td>
</tr>
<tr>
<td>c. {{speak} to} [yesterday]</td>
<td></td>
<td></td>
<td>1</td>
<td>1 W</td>
<td>1 W</td>
<td>1 W</td>
</tr>
<tr>
<td>c. [speak] [to] [yesterday]</td>
<td></td>
<td></td>
<td>2 W</td>
<td></td>
<td>1 W</td>
<td>1 W</td>
</tr>
</tbody>
</table>

At step 4, there is a critical conflict among several constraints. The winner in (63a) violates both ALIGN-R(PPh, PWd) and EXHAUSTIVITY(PPh). The losers satisfy both of these constraints, but at the expense of violating higher-ranking ALIGN-R(Lex_{max}, PPh). The ranking of ALIGN-R(Lex_{max}, PPh) above ALIGN-R(PPh, PWd) required in this tableau is the second of the two rankings that are necessary in the HS analysis but unnecessary in (though consistent with) Selkirk’s P-OT analysis.

Step 4 of HS analysis of *(Who did John) speak to yesterday?*

<table>
<thead>
<tr>
<th>[speak] to {{yesterday}}</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex_{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PWd CON</th>
<th>Exh (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → {{speak} to} {{yesterday}}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>b. [speak] to {{yesterday}}</td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td>c. [speak] [to] {{yesterday}}</td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
<td>L</td>
</tr>
</tbody>
</table>

Step 5 is the point in the derivation where the HS analysis shows that it is capable of producing top-down effects. The winner of step 4 has a phonological phrase, */[speak] to*, that does not end in a prosodic word. When prosodic words were first built at the earliest steps of the derivation, there was no reason to create a prosodic word */to*, and there was a very good reason not to: PWdCON. At that derivational stage, there was no phonological phrase structure yet, and so the winning candidates vacuously satisfied ALIGN-R(PPh, PWd). But now, at step 5,
there is phonological phrase structure present and the markedness of \{speak\ to\} is apparent. Because ALIGN-R(PPh, PWd) dominates PWdCON, the prosodic structure of to is revised:

\[(64)\] Step 5 of HS analysis of (Who did John) speak to yesterday?

<table>
<thead>
<tr>
<th>{speak to} {yesterday}</th>
<th>AL-L (LexWd, PWd)</th>
<th>AL-R (LexWd, PWd)</th>
<th>AL-R (Lex\textsuperscript{max}, PPh)</th>
<th>AL-R (PPh, PWd)</th>
<th>PWdCON</th>
<th>EXH (PPh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. {speak\ [to] {yesterday}]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. {speak\ to} {yesterday}]</td>
<td></td>
<td></td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
<td></td>
</tr>
</tbody>
</table>

The derivation converges at step 6.

The interaction in tableau (64) is a top-down effect. The presence of the phonological phrase is forcing the creation of prosodic word where none was required previously. This derivation is certainly not strictly bottom-up, and it therefore constitutes a demonstration that HS does not entail strict bottom-upness. In a strictly bottom-up derivation, decisions about the parse into prosodic words would be made once and for all before any phonological phrase is constructed. One might imagine constructing a derivational theory of prosodic structure that is strictly bottom-up in this sense, but HS is not such a theory because it has full availability of structural operations. With all structural operations available throughout the derivation, as long as there are markedness constraints that mention two levels of the prosodic hierarchy — such as ALIGN-R(PPh, PWd) — there can be top-down as well as bottom-up effects.

McCarthy’s (2002) argument for P-OT, based on Selkirk’s (1995) analysis, is really an argument against strict bottom-up derivations. But, as we have shown, this is a straw man: HS is not strictly bottom-up, and indeed it is capable of analyzing the English function word data with exactly the same constraints that mention two levels of the prosodic hierarchy — such as ALIGN-R(PPh, PWd) — there can be top-down as well as bottom-up effects.

Pater (2000: 248-250) presents a P-OT analysis of the sonorant destressing phenomenon in English, and he argues that it too demonstrates the need for parallelism. To keep the discussion manageable, we focus on a representative subset of the data: words of the form \#H\textsuperscript{X\textperiodcentered}…, where H denotes a heavy initial syllable and X is the syllable whose stressedness is the point of the analysis.

When X has an obstruent coda, it is regularly stressed:

\[(65)\] \#H\textsuperscript{X\textperiodcentered}… words, X = CVO

| timb\textsuperscript{ak’tuː} | Timbuctoo |
| di\textsuperscript{lek’tɛːfn} | delectation |
| sm\textsuperscript{teq’mɛərɛk} | syntagmatic |
| m\textsuperscript{dig’neʃn} | indignation |

When X would be expected to have a sonorant coda, however, it is unstressed and the would-be coda is parsed as the nucleus:

\[(66)\] \#H\textsuperscript{X\textperiodcentered}… words, X = CN

| g\textsuperscript{aʊŋ’zoːlə} | gorgonzola |
| mo\textsuperscript{zɪm’bɪk} | Mozambique |
| sj\textsuperscript{ɪm’tɛnɪjəs} | simultaneous |
| p\textsuperscript{ɛns’vɛnɪjə} | Pennsylvania |

Neither generalization is exceptionless (for which see Burzio 2007, Elfner 2007, Ross 1972), but we follow Pater in assuming that this is a grammatically controlled pattern that must be accounted for in any analysis.
In Pater’s P-OT analysis, the behavior exemplified in (66) is a consequence of two constraints. One, *CLASH-HEAD, rules out *ˌɡɔɹˌɡɑnˈzoːlə because of the stress clash between gan and the main-stressed syllable zoː. The other constraint is, in our terms, WSP-WORD, which is violated by *ˌɡɔɹˌɡɑnˈzoːlə because gan is a heavy syllable that is unstressed and unfooted (cf. (23)). Both of these constraints dominate *NUC/SON, which is violated by the sonorant consonant nucleus of the syllable gn. Tableau (67) presents this ranking argument.

(67) **gorgonzola** in Pater’s (2000) P-OT analysis

<table>
<thead>
<tr>
<th></th>
<th>/gɔɹɡɑnˈzoːlə/</th>
<th>*CLASH-HEAD</th>
<th>WSP-WORD</th>
<th>*NUC/SON</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>→ ˌɡɔɹɡn̩ˈzoːlə</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>ˌɡɔɹˌɡɑnˈzoːlə</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>ˌɡɔɹɡɑnˈzoːlə</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

The examples in (65) work differently, however, because *CLASH-HEAD is itself crucially dominated by *NUC/OBST, which ensures that /ˈtɪmbʌkˈtuː/ is not a viable option:

(68) **Timbuctoo** in Pater’s (2000) P-OT analysis

<table>
<thead>
<tr>
<th></th>
<th>/ˈtɪmbʌktuː/</th>
<th>*NUC/OBST</th>
<th>WSP-WORD</th>
<th>*CLASH-HEAD</th>
<th>*NUC/SON</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>→ ˌtɪmˌbʌkˈtuː</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>ˌtɪmbkˈtuː:</td>
<td>1 W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>ˌtɪmbʌkˈtuː:</td>
<td>1 W</td>
<td>L</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The argument for parallelism goes like this (Pater 2000: 250):

In this treatment of the asymmetry between sonorant-final and obstruent-final syllables, it appears to be crucial that syllabification and stress assignment be evaluated in parallel, rather than established and evaluated in sequence. Whether a syllable in a pretonic sequence is unstressed depends in part upon whether the syllable-final consonant can be parsed as a nucleus. Whether a sonorant is parsed as a nucleus in turn depends upon whether it is unstressed. This sort of interdependence between the well-formedness of stress and syllable structure is awkward to express in a theory in which syllabification derivationally precedes stress placement...

Like the argument for parallelism discussed in the previous section, this is an argument against a particular kind of serial theory and not against serial theories generally. It is an argument against strict bottom-up serialism, in which all syllabification operations precede all stress-assignment operations. Bottom-up serialism allows stress to depend on syllabification, but it does not allow syllabification to depend on stress. It is not possible to construct a strictly bottom-up serial analysis that mimics the analysis of gorgonzola in (67).

But HS is not bottom-up serialism. In HS, syllabification and stress assignment operations are available at every step of the derivation. If ˌɡɔɹɡɑnˈzoːlə is the input to some intermediate step of the derivation, then GEN can produce the candidates ˌɡɔɹɡnˈzoːlə and ˌɡɔɹˌɡɑnˈzoːlə, one of which has changed the syllabification and the other of which has changed the stress. This means that HS can in principle have exactly the same candidate competition as the P-OT analyses in (67) and (68). Thus, there is no prima facie case here for P-OT over HS.

7 **Limits on cross-level interactions in Harmonic Serialism**

P-OT and HS analyses of the same data may place different demands on **CON**. Because HS
maps underlying to surface forms via a gradual, harmonically improving derivation, it will sometimes require more finely differentiated constraints than P-OT. Because any theory of Con has typological consequences, this point of difference offers a way of falsifying a HS analysis using techniques that are standard in OT (see, e.g., McCarthy 2008: chapter 5).

To make this concrete, consider the interaction of apocope and vowel shortening in Yawelmani (Archangeli 1984, Kisseberth 1970, Kuroda 1967, Newman 1944, and many others). Yawelmani has a process of i-epenthesis that breaks up unsyllabifiable consonant clusters:

(69) Epenthesis in Yawelmani

\[
/\text{ʔilk}-\text{hin}/ \quad \text{ʔi.lik.hin} \quad \text{‘sing (nonfuture)’}
\]
\[
/lihm-hin/ \quad \text{li.him.hin} \quad \text{‘run (nonfuture)’}
\]
\[
\text{cf.} /\text{ʔilk}-\text{al}/ \quad \text{ʔil.kal} \quad \text{‘sing (dubitative)’}
\]
\[
/lihm-al/ \quad \text{lih.mal} \quad \text{‘run (dubitative)’}
\]

It also has a process that shortens long vowels in closed syllables:

(70) Closed syllable shortening in Yawelmani

\[
/lan-hin/ \quad \text{lan.hin} \quad \text{‘hear (nonfuture)’}
\]
\[
/\text{s̺ap}-\text{hin}/ \quad \text{s̺ap.hin} \quad \text{‘burn (nonfuture)’}
\]
\[
\text{cf.} /\text{lan}-\text{al}/ \quad \text{laː.nal} \quad \text{‘hear (dubitative)’}
\]
\[
/\text{s̺ap}-\text{al}/ \quad \text{s̺aː.pal} \quad \text{‘burn (dubitative)’}
\]

Together, these processes ensure that Yawelmani has no surface CVCC or CVːC syllables.

A process of apocope deletes final short vowels. When apocope would create a final CVːC syllable, vowel shortening occurs. We might also expect to see epenthesis called on when apocope would create a final CVCC syllable. In fact, though apocope is blocked in that situation.

(71) Apocop e in Yawelmani

\[
/\text{taxaː}-\text{kˀa}/ \quad \text{ta.xakˀ} \quad \text{‘bring!’}
\]
\[
/\text{taxaː}-\text{mi}/ \quad \text{ta.xam} \quad \text{‘having brought’}
\]
\[
\text{cf.} /\text{xat}-\text{kˀa}/ \quad \text{xat.kˀa} \quad \text{‘eat!’}
\]
\[
/\text{xat}-\text{mi}/ \quad \text{xa.tikˀ} \quad \text{‘having eaten’}
\]
\[
\text{not} /\text{xat}-\text{mi}/ \quad \text{xa.tim}
\]

In P-OT, the difference in (71) can be attributed to a difference in the ranking of two faithfulness constraints relative to the markedness constraint that demands apocope, *V#. *V# dominates Max-µ to account for \text{ta.xak}^2 (tableau (72)), but it must itself be dominated by Dep-V to account for *xa.tik’ (tableau (73)). In both cases, undominated *Superheavy rules out the CV:C and CVCC syllables that would result from apocope alone:
(72) \[ *V# \gg \text{MAX}(\mu) \text{ in Yawelmani (P-OT analysis)} \]

<table>
<thead>
<tr>
<th>/taxaː-k^2a/</th>
<th>*SUPER</th>
<th>*V#</th>
<th>MAX-\mu</th>
<th>MAX-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → ta.xak^2</td>
<td></td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. ta.xa:k^2a</td>
<td>1 W</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c. ta.xak^2</td>
<td>1 W</td>
<td>1 L</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

(73) \[ \text{DEP-V} \gg *V# \text{ in Yawelmani (P-OT analysis)} \]

<table>
<thead>
<tr>
<th>/xat-k^2a/</th>
<th>*SUPER</th>
<th>DEP-V</th>
<th>*V#</th>
<th>MAX-\mu</th>
<th>MAX-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → xat.k^2a</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. xatk^2</td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
<td>1 W</td>
<td></td>
</tr>
<tr>
<td>c. xa.tik^2</td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
<td>1 W</td>
<td></td>
</tr>
</tbody>
</table>

This constraint set is insufficient for a HS analysis of Yawelmani, however. The HS factorial typology of this constraint set over the inputs in (72) and (73) was computed using OT-Help (Staubs et al. 2010), under the assumption that deletion, shortening, and epenthesis each require their own derivational step (though resyllabification does not). This typology includes only the languages in (74), and Yawelmani is not among them.

(74) \[ \text{A factorial typology in HS} \]

<table>
<thead>
<tr>
<th>taxa:k^2a</th>
<th>xatk^2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>taxa:k^2</td>
<td>xatk^2</td>
</tr>
<tr>
<td>taxak^2</td>
<td>xatk^2</td>
</tr>
<tr>
<td>taxak^2</td>
<td>xatik^2</td>
</tr>
<tr>
<td>taxak^2</td>
<td>xatik^2</td>
</tr>
</tbody>
</table>

The reason why Yawelmani is not in this typology is that \(*\text{SUPERHEAVY}\), as the only constraint against both CVCC and CV:C syllables, is simply too general. At the first step of the derivation of /taxaː-k^2a/ (after initial syllabification), apocope needs to be allowed, yielding \(ta.xak^2\), which will undergo vowel shortening at the next step. At the first step of the derivation of /xat-k^2a/, however, apocope needs to be blocked, so the derivation immediately converges on faithful \(xat.k^2a\). The constraint \(*\text{SUPERHEAVY}\) is obviously inadequate to this task: if ranked above \(*V#\), it will wrongly block apocope in both \(ta.xa:k^2a\) and \(xat.k^2a\), and if ranked below \(*V#\), it will wrongly allow apocope in both \(ta.xak^2\) (\(\rightarrow \text{ta.xak}^2\)) and \(xatk^2\). At the first step of the derivation, then, there is no way of knowing that the mapping /taxaː-k^2a/ \(\rightarrow ta.xak^2\) should be allowed because at the next step the violation of \(*\text{SUPERHEAVY}\) will be eliminated by shortening, while the mapping /xat-k^2a/ \(\rightarrow xatk^2\) should be blocked by the same constraint.

This problem arises because positing a \(\text{CON}\) with just \(*\text{SUPERHEAVY}\) entails that CV:C and CVCC syllables are always equally marked. Now suppose we include \(*\text{COMPLEX-CODA}\) in \(\text{CON}\). This constraint introduces an additional bias against CVCC syllables, and thereby eliminates any need for doing shortening and apocope in parallel. At step 1, shown in tableaux (75) and (76), \(*V#\) compels violation of low-ranking \(*\text{SUPERHEAVY}\), but it is unable to compel violation of high-ranking \(*\text{COMPLEX-CODA}\).

(75) \[ \text{Step 1 of }/taxaː-k^2a/ \rightarrow ta.xak^2 \rightarrow ta.xak^2 \text{ (HS analysis)} \]

<table>
<thead>
<tr>
<th>/taxaː-k^2a/</th>
<th>*\text{COMP-CODA}</th>
<th>*V#</th>
<th>*\text{SUPER}</th>
<th>MAX-V</th>
<th>MAX-\mu</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → ta.xak^2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. ta.xa:k^2a</td>
<td>1 W</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>
Step 1 of /xat-kˀa/ → xat.kˀa (HS analysis)

<table>
<thead>
<tr>
<th>/xat-kˀa/</th>
<th>*COMP-CODA</th>
<th>*V#</th>
<th>*SUPER</th>
<th>MAX-V</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. xat.kˀa</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. xatkˀ</td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
<td>1 W</td>
<td>1 W</td>
</tr>
</tbody>
</table>

The derivation of xat.kˀa has already converged. The derivation of ta.xakˀ continues for one more step before the convergence step:

Step 2 of /taxaː-kˀa/ → ta.xaːkˀ → ta.xakˀ (HS analysis)

<table>
<thead>
<tr>
<th>ta.xaːkˀ</th>
<th>*COMP-CODA</th>
<th>*V#</th>
<th>*SUPER</th>
<th>MAX-V</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ta.xakˀ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ta.xaːkˀ</td>
<td>1 W</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The upshot is that the derivations of ta.xakˀ and xat.kˀa can be distinguished by the markedness of the intermediate form that would result from apocope: ta.xakˀ is allowed because *SUPERHEAVY is ranked below *V#, but xatkˀ is not because *COMPLEX-CODA is ranked above *V#. As we have so often seen in HS analyses, a surface-unviolated constraint (here *SUPERHEAVY) is violated at an intermediate step of the derivation of ta.xakˀ.

We have shown that the HS analysis of Yawelmani requires both *COMPLEX-CODA and *SUPERHEAVY, while the P-OT analysis requires only *SUPERHEAVY. Is this merely an unwelcome expedient to which HS is forced by its assumptions, or does even P-OT need both of these constraints? Data from Sudanese Arabic show that both constraints are indeed required in both the parallel and serial theories. This language has a syncope process that can produce surface CVC or CVːC syllables (78a), but is blocked when a CVCC syllable would result (78b).

Syncope in Sudanese Arabic (Hamid 1984: 82ff.)

<table>
<thead>
<tr>
<th>/fihim-u/</th>
<th>'fihmu</th>
<th>‘understood (m. pl.)’</th>
</tr>
</thead>
<tbody>
<tr>
<td>/jikatib-u/</td>
<td>ji'kaxdbu</td>
<td>‘correspond (m. pl.)’</td>
</tr>
<tr>
<td>/maːsik-a/</td>
<td>'maska</td>
<td>‘holding (f. sg.)’</td>
</tr>
<tr>
<td>/jaktub-u/</td>
<td>jaktibu</td>
<td>‘write (m. pl.)’</td>
</tr>
<tr>
<td>/jitardʒim-u/</td>
<td>ji'tardʒimu</td>
<td>‘translate (m. pl.)’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*jaɡdbu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*ji’tardʒmu</td>
</tr>
</tbody>
</table>

It is impossible in P-OT or HS to analyze this language without a markedness constraint that disfavors *jaɡdbu but not ji’kaxdbu. That constraint is *COMPLEX-CODA, which is therefore indispensable even in P-OT.

In OT, constraints are hypotheses about Con that are subject to empirical verification through language typology. Typology therefore provides a check on constraint proliferation. As Yawelmani illustrates, harmonic improvement in HS derivations may require a richer constraint set than a P-OT analysis of the same facts. But language typology offers a straightforward way of testing proposed constraints, with results that could in principle falsify a proposed HS analysis and even undermine the theory in which it is embedded.

As in P-OT, the logic of language typology can be used in HS to discriminate among competing hypotheses about how some data should be analyzed. McCarthy & Prince (1993b) present an analysis of word-final vowel shortening in Axininca Campa that appears to require parallelism. A HS reanalysis is possible, but not the most obvious one.

Stress in Axininca Campa is generally quantity-sensitive. In sequences of light syllables
the pattern is left-to-right iambic; every other syllable beginning with the second is stressed (examples in (79a)), and this pattern is restarted after a stress-attracting heavy syllable. Stress nearly always avoids the final syllable, however. Disyllables generally receive initial stress, as in (79b), while polysyllabic words ending in an even number of light syllables vary between shifting the last stress to the penult and omitting the stress altogether, (79c). McCarthy and Prince (1993b) analyze the penult stress cases with a final trochee.


a. tʃʰoˈrina  ‘species of palm’
   iʃʰikaˈkina  ‘he has cut me’

b. ʃarɪ  ‘macaw’
   ʃ Proto  ‘shrimp’

c. kiˈmiˈtaka ~ kiˈmitaka  ‘perhaps’
   hoˈtiˈtana ~ hoˈtitana  ‘he let me in’

Because the distinction between main and secondary stress is unimportant for our purposes here, we have not indicated it in the transcriptions.

On the surface, syllables with a long vowel or diphthong always receive stress, but this requirement is satisfied somewhat heterogeneously when it conflicts with a constraint prohibiting stress on the final syllable. When the final syllable underlyingly contains a long vowel, it surfaces as short and unstressed, as the examples in (80) show. But when the final syllable contains a diphthong, it surfaces faithfully and receives stress.

(80) Final long vowels shorten (Payne 1981: 119-121)

/sampa:/  ‘sampa ‘balsa’  cf. no-sampaː-ti ‘my balsa’
/sawo:/  ‘sawo ‘case’  cf. no-sawoː-ti ‘my case’
/cʰimi/  ‘cʰimi ‘ant’  cf. no-cʰimiː-ti ‘my ant’

(81) Final diphthongs stressed (Payne, Payne and Santos 1982: 187-188)

noˈpai  ‘my sugarcane’
ampɔˈkai  ‘we will come back’
kiˈtiʃiˈtakoˈtai  ‘we came in the morning’
ˈaːˈtai  ‘we will go’

McCarthy & Prince’s (1993b) P-OT analysis accounts for the differential treatment of final long vowels and diphthongs in terms of differing degrees of unfaithfulness. WSP, which requires all heavy syllables to be stressed, and NON-FINALITY, which requires that final syllables not be stressed, are in conflict when a diphthong or a long vowel is in the final syllable. In the long vowel case (tableau (82)), both markedness constraints are satisfied by shortening the vowel (i.e., deleting a mora), so WSP and NON-FINALITY dominate MAX-μ. In the diphthong case (tableau (83)), a violation of MAX-V would be needed for simultaneous satisfaction of the markedness constraints. Instead, we actually find that a violation of NON-FINALITY is tolerated in order to avoid violation of MAX-V and WSP. Thus, WSP and MAX-V both dominate NON-FINALITY.
Final long vowels shortened in P-OT

<table>
<thead>
<tr>
<th></th>
<th>WSP</th>
<th>MAX-V</th>
<th>NON-FIN</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>('sawo)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>b.</td>
<td>('sawo)</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>(sa'wo)</td>
<td></td>
<td>1 W</td>
<td>L</td>
</tr>
</tbody>
</table>

Final diphthongs stressed in P-OT

<table>
<thead>
<tr>
<th></th>
<th>WSP</th>
<th>MAX-V</th>
<th>NON-FIN</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>(no'pai)</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>('nopa)</td>
<td>1 W</td>
<td>L</td>
<td>1 W</td>
</tr>
<tr>
<td>c.</td>
<td>('nopai)</td>
<td>1 W</td>
<td></td>
<td>L</td>
</tr>
</tbody>
</table>

This is a CLI between a moraic process (vowel shortening) and stress. When a final heavy syllable can be shortened, stress will surface on a different syllable or be omitted, but when a final heavy syllable cannot be shortened because it contains a diphthong, it must surface with the stress. The ranking of NON-FINALITY above MAX-μ and below MAX-V determines how WSP will be satisfied in final syllables: by shortening if possible, but otherwise by stressing.

When we try to restate this analysis derivationally, however, we run into a problem. Under the ranking in (82) and (83), the first step of the derivation from /sawo:/ will produce the output (sa'wo). This is shown in tableau (84).

Step 1 from /sawo:/

<table>
<thead>
<tr>
<th></th>
<th>WSP</th>
<th>MAX-V</th>
<th>NON-FIN</th>
<th>MAX-μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. →</td>
<td>(sa'wo)</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>('sawo)</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

Throughout, we have assumed that GEN cannot shorten a vowel and move a stress together, in a single step. It follows, then, that the next step after (sa'wo) will include a candidate with shortening (sa'wo) and a candidate with stress shift ('sawo), but not a candidate with both ('sawo). Unfortunately, neither (sa'wo) nor ('sawo) is more harmonic than (sa'wo) — the former merely adds a violation of MAX-μ, and the latter swaps a violation of NON-FINALITY for a violation of higher-ranking WSP. Therefore, the derivation will immediately converge on the wrong surface form, * (sa'wo).

One obvious but typologically unsound line of attack is to separate WSP into two constraints, one requiring that long vowels be stressed and the other requiring that diphthongs be stressed. If WSP-DIPHTHONG is ranked above NON-FINALITY, then /nopai/ will correctly map to no(ˈpai) at step 1:
(85) Step 1 of /nopai/ with WSP-DIPHTHONG and WSP-V:

<table>
<thead>
<tr>
<th></th>
<th>WSP-DIPH</th>
<th>NON-FIN</th>
<th>WSP-V:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (noˈpai)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. → (ˈnopai)</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

And if WSP-V: is ranked below NON-FINALITY, then /sawoː/ will map to (ˈsawoː) at step 1, which will change to (ˈsawo) at step 2 before converging.

(86) Step 1 of /sawoː/ with WSP-DIPHTHONG and WSP-V:

<table>
<thead>
<tr>
<th></th>
<th>WSP-DIPH</th>
<th>NON-FIN</th>
<th>WSP-V:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. → (ˈsawoː)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. → (saˈwoː)</td>
<td>1 W</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

Although this analysis initially looks promising, it is typologically shaky. Two languages are known to exist in which long vowels and diphthongs have different weight in stress assignment, and in both cases the long vowels are treated as heavier than the diphthongs. (The languages are Kara and Maori (de Lacy 1997, Gordon 2006: 123).) If this is a valid typological generalization, then it must mean that WSP-DIPHTHONG never dominates WSP-V:. The ranking in (85) and (86) is inconsistent with that.

Although typological considerations militate against one imaginable account of the Axininca facts, they also point the way toward a better one. Buckley (1998) presents extensive cross-linguistic evidence that NON-FINALITY cannot explain all cases where final long vowels are prohibited, so a constraint against final length tout court is needed in CON.23 For example, Italian regularly lengthens stressed vowels in open syllables (kapiˈtaːno ‘captain’), but not if the stressed syllable is word-final (kafˈfe ‘coffee’). The dispreference for word-final long vowels extends even to the sandhi process raddoppiamento sintattico, in which lengthening of final stressed syllables is accomplished by consonant gemination rather than vowel lengthening: kafˈfeˈnɛrə ‘black coffee’.

If CON includes a constraint against final long vowels, then the raison d’être for McCarthy & Prince’s analysis vanishes. The point of their analysis is to use NON-FINALITY and the faithfulness difference between shortening a long vowel and simplifying a diphthong to explain why final long vowels are unstressed and short while final diphthongs are stressed and unchanged. This explanation is unnecessary if final long vowels are marked regardless of stress.

As the Axininca Campa and Arabic examples have illustrated, analyses of CLIs in HS may require finer differentiation of markedness constraints than analyses of the same data in P-OT. But these richer constraint sets (and ultimately HS itself) are falsifiable using techniques of typological analysis that are standard in OT (see, e.g., McCarthy 2008: chapter 5), and that have now been automated for HS with the OT-Help software package (Staubs et al. 2010).

8 Conclusion

Cross-level interactions are a good place to look for differences between parallel and serial
theories. It is no surprise that they have figured in almost all arguments for parallelism, nor that they present some of the severest challenges to rule + constraint theories.

Cross-level interactions have certain properties that can challenge derivational theories of constraint satisfaction. In many cases, they require surface-true constraints to be violated in the course of the derivation. They also show that derivations cannot always build structure in a strictly monotonic, bottom-up fashion.

We have argued that Harmonic Serialism has these characteristics. Because it is simply a derivational version of Optimality Theory, it has violable constraints. There is nothing in Harmonic Serialism that requires surface-unviolated constraints to be unviolated throughout the derivation; indeed, it is hard to imagine how this requirement could even be stipulated within the theory. Further, there is nothing in Harmonic Serialism that requires structure to be built from the bottom up; rather, every structure-building or -altering operation in GEN is available in principle at every step of the derivation.

Perhaps the most important result to emerge from this study is this. Serialism is nothing more than a vague description of a large class of theories that posit intermediate derivational steps. Critiques of serialism necessarily deal with specifics, and those specifics are not something that is known a priori. Different serial theories have very different properties, and facts that are a problem for one are very likely not a problem for all. The arguments against serialism in the OT literature are, as we have shown, arguments against specific serial theories that have very little to do with Harmonic Serialism.

Notes

1 This research was supported by grant BCS-0813829 from the National Science Foundation to the University of Massachusetts Amherst. We are indebted to all of the participants in our weekly grant group for their advice about this chapter, with particular thanks going to Matt Wolf for reviewing the manuscript.

2 *(aˈʧo)(ˈwo)<wo> also violates *CLASH, but that cannot be the reason why it is ill-formed, because Hixkaryana is otherwise quite tolerant of clash: *(ˈnak)(ˈɲoh)(ˈyatʃ)(keˈnaː)no ‘they were burning it’.

3 For copious references to previous work on SWP, see Gouskova (2003: 90fn.).

4 Kisseberth (1970: 305) does not claim to have solved the conspiracy problem — he offers his proposal about the formalization of (non-)blocking as a “first tentative step in the construction of a theory of phonology employing the notion of derivational constraints.”

5 Calabrese (2009) proposes a constraint + repair system that appears to be capable of accounting for both blocking and triggering effects involving markedness constraints that are surface-inviolable. Each active markedness constraint is associated with a ranked list of repairs, but markedness constraints themselves are not ranked. If the input to the derivation violates some markedness constraint M1, its first repair R1M1 is triggered. If the result of applying R1M1 violates M2, then its first repair R1M2 is triggered, and so on. The derivation terminates when all markedness constraints have been satisfied. If the derivation reaches a point where a markedness constraint MB is violated and there is no viable repair, then a kind of blocking occurs: the derivation crashes and a new derivation begins in which M1’s second-ranked repair R2M1 is triggered.

The biggest problem with this theory is its commitment to surface-inviolability of active markedness constraints. As a result, FOOT-BINARITY cannot block foot parsing in Hixkaryana (aˈʧo)wo<wo>, though it can trigger lengthening in *(ˈkaː)<na>. The problem is that the
standard foot-parsing markedness constraint, PARSE-SYLLABLE ("assign one violation mark for every unfooted syllable") is not surface-true in Hixkaryana, so it could not be active, according to this theory. That problem could be circumvented if PARSE-SYLLABLE were replaced by a constraint against a sequence of two unfooted syllables, PARSE-2 (Kager 1994), which is surface-unviolated in Hixkaryana. But this move leads to an obvious loss of generality: there are now two foot binarity constraints, FOOT-BINARITY itself and PARSE-2. This loss of generality is not at all atypical of efforts to recode violable constraints as inviolable (McCarthy 2002: 16-17).

6 The constraint SWP uniformly favors heavy stressed syllables in both iambic and trochaic feet. This predicts that stressed syllable lengthening should not be unique to iambic languages. Although Hayes (1995: 82ff.) argues that stressed syllable lengthening is most robustly attested in iambic languages, this generalization does not appear to be categorical. For discussion of trochaic lengthening, see for example Hyde (2007), McGarrity (2003), and Revithiadou (2004).

7 Kager’s analysis of Hixkaryana will not work if UNEVEN-IAMB is replaced with the better-motivated constraint SWP. The problem is clear from tableau (11): if UNEVEN-IAMB were replaced by SWP, then *(aˈʧoː)(ˈwoː)wo in (11b) would wrongly win. The HS analysis does not have this liability — see section 4 for further discussion.

8 The prosodic hierarchy (Selkirk 1980) imposes a limited amount of bottom-upness. For example, it is a standard assumption in prosodic hierarchy theory that feet are constructed out of syllables and that feet cannot have unfilled daughter nodes. If this is taken to be an inviolable constraint, then no foot can be built until at least one syllable has been.

9 Iambic shortening also affects words with final closed syllables: (ˈputat) can be scanned in Plautine verse as LL rather than LH.

10 Mester’s analysis is not without its critics; see Fortson (2008) and Lahiri et al. (1999).

11 Prince & Smolensky (1993/2004: 62) have a single non-finality constraint that is violated once for a word-final main-stressed foot or syllable, and twice for both. In keeping with the standard practice in later work, we have split it into two constraints.

12 Candidate (27d) also violates PARSE-SYLLABLE. That is not the reason why it loses to (ˈa.mot), however, because NON-FINALITY(ˈft) dominates PARSE-SYLLABLE, as shown by (29e).

13 The P-OT and HS analyses make different predictions about foot parsing in syllables that precede the main stress. For example, when the main stress is preceded by two heavy syllables, as in audiːˌtotte ‘hear! (fut. pl.)’, the P-OT analysis predicts (au)(di)(ˈtoː)te while the HS analysis predicts (audiː)(ˈtoː)te. Little or nothing is known for sure about pretonic secondary stress in Latin, and existing conjectures (e.g., initial secondary stress if main stress is on the perinital (Allen 1973: 190-191)) are unable to settle empirical questions of such subtlety.

14 They credit a 1992 manuscript by Robert Kirchner.

15 The quality of the epenthetic vowel is a result of emergence of the unmarked (McCarthy and Prince 1994). Epenthetic a maximizes sonority in the nucleus, as required by the peak hierarchy (Prince and Smolensky 1993/2004: 151ff.).

16 DEP(place) or its equivalent is also needed in any P-OT analysis. Flap-final /ter/ augments as ˈte.ɾa, not ˈter.ta, because of an undominated constraint against r+coronal clusters. DEP(place) is needed to rule out *ter.pa and *ter.ka, which are phonotactically impeccable.

17 The change of /t/ is optional, according to Klokeid (1976: 39). There is variation between t and r word-finally, according to Hale & Nash (1997).

18 Although ALIGN-R(MWORD, σ) dominates *CODA(lab/dors), it must not prevent
*CODA*(lab/dors) from favoring deletion of \( \eta \) in /\text{wuŋkun}\_\eta/ → *wuŋkunu*. Prince & Smolensky’s (1993/2004: 127) definition of ALIGN-R(MWORD, \( \sigma \)) reads like this: “The final edge of a Morphological Word corresponds to the final edge of a syllable.” Under this definition, ALIGN-R(MWORD, \( \sigma \)) would wrongly block deletion in /\text{wuŋkun}\_\eta/ as well as correctly trigger epenthesis in /\text{kaŋ}/. In correspondence theory (McCarthy and Prince 1995, 1999), ANCHOR constraints decouple these two senses of alignment. One type of ANCHOR constraint says that the underlying word-final consonant must have an output correspondent. This constraint is dominated by *CODA*(lab/dors), as shown by /\text{wuŋkun}\_\eta/ → *wuŋkunu*. The other type of ANCHOR constraint says that if there is an output correspondent of an underlying word-final consonant, then it must be syllable-final. This is the sense of ANCHOR that we impute to ALIGN-R(MWORD, \( \sigma \)) in the analysis of Lardil.

19 We arbitrarily assume that the underlying vowel of the medial syllable of *gorgonzola* is /\text{a}/.

20 This language is now known as Ajy’ininka Apurucayali.

21 The system is more complicated than heavy/light, however. Payne et al. (1982) propose a weight/prominence continuum to account for the relative likelihood of syllable types attracting or repelling stress.

22 Examples like *sima/no-sima-ni* ‘fish’/‘my fish’ show that this process is indeed shortening rather than lengthening before a suffix. Examples like *mi*: ‘otter’ show that shortening is blocked in monosyllables — a typical word-minimality effect (McCarthy and Prince 1993b: 164).

23 On the functional basis of final vowel length neutralization, see Myers & Hansen (2007).

24 CLIs are not the only source of arguments for parallelism in the OT literature. One other prominent argument is based on reduplicative overapplication (McCarthy and Prince 1995, 1999); for a HS approach, see McCarthy, Kimper, and Mullin (2012); for a general critique of the argument, see among others Inkelas and Zoll (2005) and Kiparsky (2010). Another prominent argument comes from phonologically-conditioned allomorphy (Mascaró 1996, Mester 1994, Tranel 1996, among others); for a HS approach, see van Oostendorp (2009) and cf. Wolf (2008); for recent discussion of the issue and references to related work, see Nevins (2011).

References


Blackwell.


Ross, John Robert (1972) A reanalysis of English word stress. In Michael Brame (ed.), Contributions to Generative Phonology, 229-323. Austin, TX: University of Texas
Press.