To appear in Toni Borowsky, Shigeto Kawahara, Takahito Shinya and Mariko Sugahara (eds.) *Prosody Matters: Essays in Honor of Elisabeth O. Selkirk*. London: Equinox Press.

Serial Harmonic Grammar and Berber Syllabification[*]

Joe Pater
University of Massachusetts, Amherst

This paper introduces serial Harmonic Grammar, a version of Optimality Theory (OT; Prince and Smolensky 1993/2004) that reverses two of Prince and Smolensky's basic architectural decisions.[1] One is their choice of constraint ranking over the numerically weighted constraints of its predecessor, Harmonic Grammar (HG; Legendre, Miyata and Smolensky 1990; see Smolensky and Legendre 2006 and Pater 2009 for overviews of subsequent work). The other is their choice of parallel evaluation over a version of OT in which the representation is changed and evaluated iteratively (Harmonic Serialism; Prince and Smolensky 1993/2004: ch. 2, McCarthy 2007 *et seq.*). I introduce serial HG with an analysis of syllabification in Imdlawn Tashlhiyt Berber (Dell and Elmedlaoui 1985, 1988, 2002), the same case that Prince and Smolensky use to introduce OT. This analysis illustrates advantages of both serialism and weighted constraints. I also discuss some of the positive consequences of the adoption of serialism for the typological predictions of HG, as well as some outstanding issues for further research on serial versions of both OT and HG.

1. Serialism and Berber syllabification

I start by showing how a serial approach to syllabification allows for an improved constraint-based analysis of Imdlawn Tashlhiyt Berber (ITB; Dell and Elmedlaoui 1985, 1988, 2002). With serial syllabification, the margin constraints of Prince and Smolensky (1993/2004: ch. 8) can be replaced by a constraint that assigns penalties to nuclei according to their position on a sonority scale. This allows the preference for high sonority nuclei in ITB to be ascribed to the same constraint that generates absolute restrictions on nuclear sonority in many languages, including a ban on obstruent nuclei in phrase-final position in ITB. The elimination of margin constraints is

---

[*] I am happy to present to Lisa Selkirk a paper that builds on her pioneering work on the role of sonority in syllabification, that discusses data from Berber, whose phonology Lisa's own work has helped us to understand, and that even manages to sneak in a few references to French. In these small ways it pays tribute to her influence on me and on the rest of the field as a scholar. I am even happier to be able to take this occasion to pay personal tribute to Lisa in her role as a colleague, mentor and friend; *merci pour tout!* This research has also especially benefited from discussion with John McCarthy and Donca Steriade, as well as with participants in Ling 730, UMass fall 2007, Ling 751, UMass spring 2010, the 2008-2009 McCarthy-Pater NSF grant group, and colloquia at University of North Carolina, University of Chicago, SUNY Stony Brook and the University of Southern California. Thanks also to Karen Jesney, Shigeto Kawahara, Jason Riggle, Brian Smith, and two anonymous reviewers for comments on an earlier draft, and to Patrick Pratt and Robert Staubs for indispensible computational help. This research was supported by grant BCS-0813829 from the National Science Foundation to the University of Massachusetts, Amherst.

[1] See also Goldsmith (1991, 1993a, 1994) and other papers collected in Goldsmith (1993b) for related precedents whose formalisms differ in various ways from the present framework. Goldsmith's Harmonic Phonology, for example, incorporates both serialism and the notion of harmonic improvement, but it does not formalize constraint interaction as in OT/HG (see Prince and Smolensky 1993/2004: 239-240, though cf. Goldsmith, 1991: fn. 10 for a brief discussion of an OT-like model).

not just a reductionist gain in elegance, since as we will see below, they are known to make unsupported typological predictions.

1.1 Background

In ITB, any consonant can occupy nuclear position. Dell and Elmedlaoui 1985 (henceforth DE85) provide the following examples of syllabicity alternations that occur across the 3rd person masculine (marked with /i-/) and the 3rd person feminine (marked with /t-/) forms of perfective verbs. Syllable edges are indicated with parentheses, and syllabic consonants are capitalized.

(1)	/i + *root*/	/t + *root*/	
	(il)(di)	(tL)(di)	'pull'
	(ir)(ba)	(tR)(ba)	'carry on one's back'
	(in)(da)	(tN)(da)	'shake (milk)
	(if)(si)	(tF)(si)	'untie'
	(ix)(si)	(tX)(si)	'go out (fire)'

Despite the relative freedom of consonants to take on the role of nuclei, ITB does show preferences for nuclei of high sonority. DE85: 108 provide the now-famous example in (2). The asterisked form is composed of *legal* syllables, but the actual syllabification has *better* syllables in that they contain only vocalic nuclei.

(2) (rat)(lult) *vs.* *(ra)(tL)(wLt) 'you will be born'

The DE85 syllabification algorithm builds core syllables consisting of an onset and a nucleus, where the nucleus has a designated sonority level. Syllables with nuclei of the highest level of sonority are built first, followed by syllables of the next highest level of sonority, and so on through the sonority scale. The sonority scale used by DE85 is as in (3).

(3) low vowel > high vowel > liquid > nasal > voiced fricative > voiceless fricative > voiced stop > voiceless stop

Once the core syllabification subroutine is finished, codas are added, and then other rules apply. The DE85 algorithm is illustrated in (4). We will return to prepausal annexation shortly.

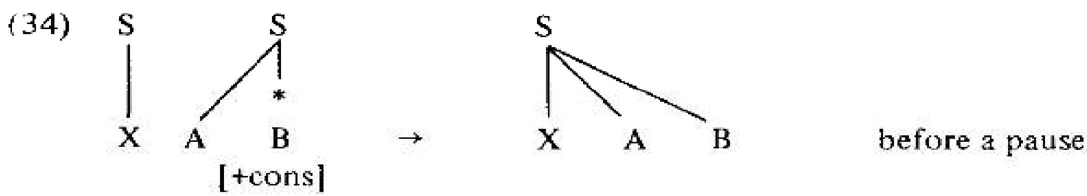
(4)	(ra)tlult	Build O-N, N = low vowel
	(ra)t(lu)lt	Build O-N, N = high vowel
	(ra)t(lu)(lT)	Build O-N, N = stop
	(rat)(lu)(lT)	Coda adjunction
	(rat)(lult)	Prepausal annexation

This analysis is an instance of an approach to preferences that can be referred to as "markedness-based rule ordering": the unmarked structure is built earlier in the derivation than the marked one, thus capturing the preference for the unmarked structure. Here, vowels are taken as nuclei before consonants, thus yielding (rat)(lult) instead of *(ra)(tL)(wLt).

A general problem with markedness-based rule ordering is that it fails to formally capture the relationship between "soft" preferences and related "hard" restrictions. Examples of hard restrictions related to the Berber preference are found in languages in which only a subset of segments can appear in nucleus position. For instance, languages like French tolerate no consonantal nuclei at all, while languages like English allow sonorants, but not obstruents, as nuclei. In these and other cases, the permissible nuclei are of higher sonority than the disallowed ones. The problem is illustrated by the fact that the early syllabification of vowel nuclei in Berber that produces the preference is in no way related to the morpheme structure constraint or filtering rule that would be used to disallow consonantal nuclei in a rule-based analysis of French.

A striking further illustration of this problem can be found in Berber itself, which has an absolute restriction against obstruent nuclei in phrase-final position. DE85's rule of prepausal annexation (p. 119) is given as follows:

(5) *Prepausal annexation*



As the derivation in (4) shows, it is this rule that is responsible for the syllabification of /ratlult/ as (rat)(lult) rather than *(rat)(lu)(lT). The issue is that the rule is simply stipulated to apply obligatorily to obstruents (see section 2.2 below); there is no connection drawn to the general relative ill-formedness of obstruents as nuclei in Berber and other languages.

Much of the success of OT comes from its ability to provide formal accounts of preferences that use the very same constraints that also yield hard restrictions.[2] As a simple example, we can take two effects of the constraint ONSET. Ranked above a competing faithfulness constraint like DEP, which penalizes epenthesis (McCarthy and Prince 1999), ONSET yields a hard restriction, which is enforced by consonant insertion, as in (6). I use numbers, rather than asterisks, to indicate the violation count.

(6)

/an/	ONSET	DEP
an	1 !	
☞ [ʔan]		1

However, even when the ranking of the two constraints is reversed, producing a language that tolerates onsetless syllables, ONSET will still prefer an intervocalic consonant to be syllabified with the following vowel, rather than with the preceding one, as shown in (7).

---

[2] The success of OT in this regard has been discussed under the rubrics of "the emergence of the unmarked", "minimal violation" and "nonuniformity". See McCarthy (2002) for discussion and references.

(7)

/an/	DEP	ONSET
☞ [an]		1
ʔan	1 !	

/aba/	DEP	ONSET
☞ [a.ba]		
ab.a		1 !

Markedness-based rule ordering usually accounts for onset preferences of this type by ordering onset creation before coda formation, as in DE85's use of early core syllable formation (Steriade 1982; see Kenstowicz 1994 and Blevins 1995 for overviews of serial approaches to syllabification). Again, the issue is that the preference is not formally related to the hard restriction, which would be put down to a separate morpheme structure constraint or epenthesis rule.

Despite this general success in formally relating preferences to absolute restrictions, there appears to be no existing OT analysis of ITB syllabification that relates its preference for high sonority nuclei to the related absolute restrictions in Berber and elsewhere. Just as the move from inviolable to minimally violable constraints opened up the general possibility of unified analyses of hard and soft restrictions, it turns out that moving from parallel to serial evaluation allows the Berber preference to be driven by the same constraint that delivers the hard restrictions.

1.2 A new constraint-based serial analysis of ITB

In this section I provide a slightly modified version of Prince and Smolensky's (1993/2004; henceforth PS93) serial analysis of ITB syllabification to illustrate how the serial version of OT works, and to provide a new argument for it. The PS93 serial analysis uses the "Nuclear Harmony Constraint".

(8) H-NUC
A higher sonority nucleus is more harmonic than one of lower sonority.

This constraint is unusual in OT because it directly states a preference. Other constraints in PS93, and almost all OT constraints since, assign violation scores. McCarthy (2003) suggests the restatement of H-NUC provided in (9), which I call *C-NUC (as we will soon see, H-NUC and *C-NUC are not fully equivalent). It assigns violation marks according to the degree to which a nucleus diverges from the ideal in terms of sonority, the maximally sonorous segment [a].

(9) *CONSONANTAL-NUCLEUS (*C-NUC)
Assign a violation mark to a nucleus for each degree of sonority separating it from [a]

McCarthy (2003) rejects *C-NUC as a parallel OT constraint for reasons to be discussed in section 2. For the Berber data to be discussed in this section, *C-NUC makes the desired choices between candidates, so long as evaluation is serial rather than parallel.

The statement of *C-NUC presupposes an *n*-ary sonority feature, as in Hankamer and Aissen (1974) and Selkirk (1984). For the rich set of sonority distinctions that DE85 document for ITB (as in (3)), we would need a correspondingly rich set of gradations in the sonority scale, so that *C-NUC would assign violations as illustrated in (10) for the high end of the scale.

(10) *C-NUC(a) = 0 *C-NUC(i) = 1 *C-NUC(r) = 2,...

To deal with the examples to be discussed in this paper, we can simplify, with *C-NUC making just the distinctions in (11).

(11) **C-NUC violations*
Vowel = 0, Sonorant Consonant = 1, Fricative = 2, Stop = 3

For their analysis, PS93 also use a constraint against hiatus to favor CV syllabification. I will use the name *HIATUS to transparently reflect how it assigns violation scores (PS93 call it ONSET):

(12) *HIATUS
Assign a violation mark to a sequence of adjacent nuclei

The interaction between Gen (the candidate creation function) and Eval (the choice function) differs in the serial version of OT from the more familiar parallel version. PS93: 21 provide the following definition of the Gen function for their serial analysis of ITB syllabification:

(13) Gen (input_i):
The set of (partial) syllabifications of input_i which differ from input_i in no more than one syllabic adjunction.

The adjunctions used here are provided in (14). I use the general term "operation" to refer to these and other transformations in Gen. DE85's core syllable formation is replaced by separate steps of Nucleus Projection and Dependent Adjunction. As John McCarthy (p.c.) points out, using core syllable formation in Harmonic Serialism would amount to building an effect of the ONSET constraint into Gen (cf. PS93).

(14) *Operations*
NUCLEUS PROJECTION
Input: Unsyllabified segment
Output: Segment syllabified as nucleus/head
e.g. a → (a) m → (M)

DEPENDENT ADJUNCTION (ONSET/CODA)
Input: Unsyllabified segment adjacent to a syllable
Output: Segment incorporated into that syllable as a dependent
e.g. t(a) → (ta) (M)t → (Mt)

In serial OT as in parallel OT, the candidates emitted by Gen are evaluated by the constraint hierarchy, which chooses the optimal candidate(s). The difference is that the optimum is resubmitted to the Gen function, and to subsequent evaluation by the constraint hierarchy. This loop continues until convergence, that is, until the unchanged candidate is picked as optimal.

The first step is illustrated in (15) for the verb [kšm] 'enter' (Dell and Elmedlaoui 1988). Here we see all of the candidates formed by the projection of a nucleus (I use Americanist [š] since it can be capitalized). I assume an undominated PARSE-SEG constraint (PS93), or feature of Gen, that forces syllabification at the expense of *C-NUC. This means that when the input contains an unsyllabified segment, the fully faithful candidate is not in the candidate set (cf. fn. 9, and section 2.4). The selection of the optimum is done as in standard OT; kš(M) performs best on *C-NUC.

(15) *Step 1*

/kšm/	*HIATUS	*C-NUC
(K)šm		3!
k(Š)m		2!
☞ kš(M)		1

The next tableau has as an input the optimal form from step 1. Only the new unshared violations are shown in the tableaux (in this case, the *C-NUC violations incurred by [M] are omitted). Here adding the onset to the previously established syllable violates no constraint, and so is optimal.

(16) *Step 2*

kš(M)	*HIATUS	*C-NUC
(K)š(M)		3!
k(Š)(M)	1!	2
☞ k(šM)		

In the third step, we're forced to put the stop in the nucleus. The candidate set for the next step would contain only the unchanged candidate, since there are no operations to apply, and the derivation would terminate.

(17) *Step 3*

k(šM)	*HIATUS	*C-NUC
☞ (K)(šM)		3

This replicates the result of the DE85 algorithm. As in DE85, the least marked nucleus is created first, followed by nuclei of decreasing sonority. The difference between DE85's rule ordering analysis and the present serial constraint-based analysis is that here the ordering, and the resulting preference, is a direct consequence of the operation of the *C-NUC constraint. The advantage of this difference is illustrated in section 2, where we see that in serial HG this same constraint also yields the related hard restrictions.

These constraints produce a different result in parallel OT, shown in (18). The fricative is chosen as a nucleus because it minimizes violations of *C-NUC.

(18) *Parallel OT*

/kšm/	*HIATUS	*C-NUC
(K)(šM)		4!
☞ (kŠm)		2

The serial OT result is correct.[3] It differs from the parallel OT one because of the lack of derivational look-ahead in the serial model. In the first step of the derivation in (15), the best available nucleus is chosen, irrespective of the eventual result of this move, which is the formation of the worst possible nucleus in the last step in (17). The derivation is blind to the fact that the globally optimal outcome would place the fricative in nucleus, as in (18). This property of derivational "myopia" (Wilson 2003) is the fundamental difference between serial and parallel versions of OT, and the basis of many of the arguments in its favor (see *esp*. McCarthy 2007a, Pruitt 2008; see further 2.5 below).

In one parallel analysis, PS93: ch. 2 use H-NUC, which positively rewards high sonority nuclei through comparison, as illustrated in (19).

(19)

/kšm/	*HIATUS	H-NUC
☞ (K)(šM)		
(kŠm)		M > Š !

While this analysis succeeds for the Berber facts, it not only uses an unusual constraint type, but it also fails to deal with absolute restrictions, as discussed in PS93: ch. 8. Their revised analysis uses a set of constraints that is in a fixed ranking and that penalize margins (onsets and codas) according to their sonority. The constraints are arranged in the fixed ranking such that a higher sonority margin always violates a higher ranked constraint than a lower sonority margin. In this case, the constraint labeled *MARGIN-NASAL dominates *MARGIN-FRICATIVE, thus picking the correct candidate.

(20)

/kšm/	*MARGIN-NASAL	*MARGIN-FRICATIVE
☞ (K)(šM)		1
(kŠm)	1!	

PS93: 191 note that by collapsing the sonority preferences for onsets and codas, the margin constraints make undesired typological predictions. They cite Prince (1983), Zec (1988, 1995) and Clements (1990) for the observation that when a language admits only a subset of consonants as codas, these are ones of relatively high, rather than low sonority. The margin constraints predict exactly the reverse. As the tableau in (20) shows, it is crucial for the analysis of ITB that the margin constraints do penalize high sonority codas. This issue has apparently never been resolved.

The point that rising sonority sequences like [kšm] require margin constraints rather than peak constraints (≈ *C-NUC) is due Donca Steriade (p.c.). PS93 and Clements (1997) focus on falling sonority sequences, which pose the same problem, as shown in tableau (21) for [ršq] 'be happy' (Dell and Elmedlaoui 1988).

[3] This is the syllabification based on evidence from poetic meter (Dell and Elmedlaoui 1988). In phrase-initial position, the syllabification based on introspection (Elmedlaoui's judgment; DE85) would have the initial stop as starting a cluster (i.e. [(kšM)]). An approach parallel to the one proposed for phrase-final position in section 2 seems possible, though I do not pursue this here.

(21) *Parallel OT*

/ršq/	*HIATUS	*C-NUC
(R)(šQ)		4!
☞ (rŠq)		2

The correct form is (R)(šQ), which we again get from serial OT with these constraints, as shown in the derivation in (22). Step 2 of this derivation shows the need to have *HIATUS and NOCODA ranked above *C-NUC to get the correct result (see also PS93: fn. 49 on the need for NOCODA >> *C-NUC for a serial OT analysis).

(22) *Serial OT Derivation for a falling sonority cluster*

a. Step 1

/ršq/	*HIATUS	NOCODA	*C-NUC
rš(Q)			3!
r(Š)q			2!
☞ (R)šq			1

b. Step 2

(R)šq	*HIATUS	NOCODA	*C-NUC
(Rš)q		1!	
(R)(Š)q	1!		2
☞ (R)š(Q)			3

c. Step 3

(R)š(Q)	*HIATUS	NOCODA	*C-NUC
(Rš)(Q)		1!	
(R)(Š)(Q)	2!		2
☞ (R)(šQ)			

As well as having typological problems (see further Clements 1997), margin constraints cannot account for the English or French restrictions on possible nuclei, since consonantal nuclei do not violate the constraint. Consider, for example, the French pronunciation of *table*, which unlike English, syllabifies the final liquid as a margin, rather than as a nucleus. I will assume that this violates a constraint against coda clusters, *COMPLEX.[4] The tableau in (23) shows that that *C-NUC correctly disfavors the English-like (tæ)(bL) *vis-à-vis* the French (tæbl). Again, I include only unshared violations of the two constraints.

(23) *French ranking*

/tæbl/	*C-NUC	*COMPLEX
(tæ)(bL)	2!	
☞ (tæbl)		1

[4] Under an analysis of French in which such final clusters are syllabified with a final empty-headed syllable (e.g. Charette 1991), the operative constraint would be one against null nuclei.

If we try to replace *C-NUC with margin constraints, French becomes impossible to generate, since (tæbl) is harmonically bounded by (tæ)(bL).

(24) *French impossible with margin constraints only*

/tæbl/	*MARGIN-LIQUID	*COMPLEX
☞ (tæ)(bL)		
(tæbl)	1	1

PS93 do not suggest that the margin constraints could replace their peak constraints, the fixed ranking equivalents of *C-NUC. This example shows why they cannot.

2. Harmonic Grammar and ITB

We have just seen that a serial OT analysis of ITB allows us to use *C-NUC, a constraint that unlike H-NUC assigns violation scores, and that unlike the margin constraints is typologically well supported. This provides an argument for serial OT, alongside those developed in McCarthy's (2007a,b, 2008a,b, 2010) recent work, and in Pruitt (2008), Elfner (2009), Jesney (to appear) and Kimper (to appear). In this section we see how switching from ranked to weighted constraints also yields benefits for the analysis of Berber syllabification (see Pater 2009 for an introduction to HG and an overview of other arguments for weighted over ranked constraints, as well as Jesney 2009 and Potts, Pater, Jesney, Bhatt and Becker 2010).

2.1 Cross-linguistic hard restrictions on nuclear sonority

A benefit of adopting weighted constraints is the resulting ability of *C-NUC to deal with the hard restrictions on nuclear sonority in languages like English and French, along with the Berber phrase-final one. Building on the discussion of H-NUC in PS93: ch. 8, McCarthy (2003) rejects *C-NUC as an OT constraint because it cannot handle inventory restrictions and other phenomena that require the constraint to impose a "cut-off" mid-way through the sonority scale. The English nucleus inventory provides an example, because it includes sonorant consonants (in unstressed syllables), but no lower sonority nuclei. If we consider the interaction of *C-NUC with DEP, we only get two languages in OT, as shown in (25) (note that with the full version of *C-NUC needed for the DE85 sonority scale, the language in 25b. would not correspond to French, since it would permit only low vowels as nuclei).

(25) a. DEP >> *C-NUC
All segments can be nuclei (Berber)

b. *C-NUC >> DEP
Only vowels can be nuclei (French)

English is thus impossible with this constraint set in OT. The picture is different in HG, as illustrated in the pair of tableaux in (26). In the HG tableaux, violations are indicated with negative integers, and the weights of the constraints are shown beneath the constraint names. The harmony of a candidate is the weighted sum of constraint scores; this number is shown at the end of each candidate's row. The candidate with the highest harmony is optimal; since the violations are negative, and the weights are positive, this will be the number closest or equal to zero. The symbol [V] stands for an epenthetic vowel. Epenthesis is worse than a sonorant consonant as a nucleus, because the weight of DEP is higher than the weight of *C-NUC. Because a fricative nucleus incurs a penalty of –2 on *C-NUC, the harmony of that candidate is lower than the epenthetic alternative.

(26) *"English" – only sonorants as consonantal nuclei*

/tn/	DEP 1.5	*C-NUC 1	
☞ (tN)		–1	–1
(tVn)	–1		–1.5

/ts/	DEP 1.5	*C-NUC 1	
(tS)		–2	–2
☞ (tVs)	–1		–1.5

The values of the weights in the tableaux in (26) are partially arbitrary, since there is an infinite number of weights that will pick the correct optima. Their non-arbitrary aspect is that they must meet the weighting conditions in (27):

(27) /tn/ → (tN), *(tVn) *w*(DEP) > *w*(*C-NUC)
/ts/ → (tVs), *(tS) 2*w*(*C-NUC) > *w*(DEP)

The weight of DEP must be greater than the weight of *C-NUC, so that a single violation of the markedness constraint is tolerated as in (tN), but it must be less than twice the weight of *C-NUC, so that epenthesis is preferred over two violations, as in (tVs) *vs.* *(tS). This example shows that a weighted version of *C-NUC can impose a cut-off mid-way through the scale, between fricatives and sonorants – a full analysis of English would account for the lack of nasal consonants in the nuclei of stressed syllables.

In Berber, DEP is weighted high enough to force any degree of *C-NUC violation. The weight of 4 will be sufficient if the maximum violation count on *C-NUC is the 3 we are assuming here (see (11)). The tableaux in (28) contrast with the English ones in that a nuclear fricative is now chosen over epenthesis.

(28) *Berber – all consonants as nuclei*

/tn/	DEP 4	*C-NUC 1	
☞ (tN)		–1	–1
(tVn)	–1		–4

/ts/	DEP 4	*C-NUC 1	
☞ (tS)		–2	–2
(tVs)	–1		–4

In French, *C-NUC is weighted higher than DEP. Here I am assuming that the constraint does not penalize vowels; if it penalized high vowels (see (10)), the constraint weights could be adjusted so as to permit them, but no segments of lower sonority.

(29) *French – no consonants as nuclei*

/tn/	*C-NUC 1.5	DEP 1	
(tN)	–1		–1.5
☞ (tVn)		–1	–1

/ts/	*C-NUC 1.5	DEP 1	
(tS)	–2		–3
☞ (tVs)		–1	–1

These typological results can be checked in OT-Help 2.0 (Staubs, Becker, Potts, Pratt, McCarthy and Pater 2010), which uses Recursive Constraint Demotion (Tesar and Smolensky 2000) to determine if sets of optima are consistent in OT, and uses an application of linear programming (Potts *et al*. 2010) to make the same determination for HG. Computational help is perhaps unnecessary to determine that HG generates just the three attested languages, while OT generates only Berber and French,[5] but this software is indispensible for comparisons between the theories in more complex cases.

The HG scalar constraint typology presented in this section, along with the serial analysis of Berber (see 2.2 for the HG version), add up to the goal set up in section 1.1: a single constraint gets absolute restrictions on the sonority of nuclei in English and French as well as the preference for highly sonorous nuclei in Berber. As far as I know, this is the first analysis to succeed in this way. The issues for the DE85 and PS93 analyses of ITB syllabification were discussed in section 1. Clements (1997) and Dell and Elmedlaoui (2002) use a SONORITY PEAK constraint for Berber that can be stated for present purposes as in (30):[6]

(30) SONORITY PEAK
If a segment is of greater sonority than the segments adjacent to it, it is syllabified as a nucleus

The Sonority Peak constraint will not generalize to the related hard restrictions. For example, it does not penalize a syllable like (tS), which is ill-formed in English and French.

2.2 Phrase-final nucleus restrictions in Berber

Because of the inadequacies of scalar constraints in OT, scalar phenomena are standardly analyzed with either a set of constraints in a fixed ranking, or with a set of constraints in a specific-to-general or stringency relation (Prince 1997, de Lacy 2004). The last section's brief demonstration of the basic adequacy of scalar constraints in HG opens a large and important topic for further research: the comparison of OT with fixed rankings or stringency relations to HG with scalar constraints (see Flemming 2001 on HG with phonetic scales). In this section, I provide a further demonstration of the utility of weighted constraints for a treatment of ITB syllabification, by showing that they allow *C-NUC to be recruited in the analysis of the sonority-based restrictions on phrase-final nuclei.

The first step is to add a general constraint penalizing final nuclei (cf. McCarthy and Prince's 1994 word-final-C constraint; see also Flack 2009 on parallels between word- and phrase-level constraints).

---

[5] OT-Help input files for these tableaux, as well as for other examples in this paper, can be found at http://people.umass.edu/pater/pater-othelp-files.zip.

[6] Here, as elsewhere in the paper, I abstract from sonority "plateaux"; see Dell and Elmedlaoui (2002) for discussion.

(31) *FINAL-N
Assign a violation mark to a phrase-final nucleus

As shown in (32), the additive interaction of *C-NUC and *FINAL-N will produce the desired result: codas, which are generally disfavored, are created phrase finally, in order to avoid a phrase-final obstruent nucleus. To focus on the crucial interaction, I omit the candidate (rat)(lu)(Lt), which would be ruled out by *HIATUS, and once more show only violations that are not shared by the candidates.

(32) *“Prepausal annexation” as a gang effect*

/ratlult/	*CODA 4	*FINAL-N 2	*C-NUC 1	
☞ (rat)(lult)	–1			–4
(rat)(lu)(lT)		–1	–3	–5

This is a “gang effect” in that the weighted sum of the violations of the two constraints *C-NUC and *FINAL-N is greater than the weighted violation of *CODA, but the weighted violations of the individual constraints is lower than that of *CODA. In our serial OT analysis, we have seen that a single violation of *CODA must trump the most severe violation of *C-NUC, the 3 incurred by a stop (see 22b; note that this tableau is for non-phrase-final position). In HG terms, this requires *CODA to have a weight greater than 3 times the weight of *C-NUC; the (4,1) weighting in (32) meets this condition.[7]

*FINAL-N is also not strong enough to beat *CODA on its own. For example, sonorant consonants can (optionally) surface as final nuclei (DE85).

(33) *Sonorants as final nuclei*

(i)(gi)(dR) ‘eagle’ (R)(gL) ‘lock’ (du)(mN) ‘they (m.) last’

This is captured in the current analysis in that a violation score of –1 on *C-NUC contributed by the sonorant is not sufficient to help *FINAL-N overcome *CODA:

(34) *Escape from prepausal annexation*

/igidr/	*CODA 4	*FINAL-N 2	*C-NUC 1	
☞ (i)(gi)(dR)		–1	–1	–3
(i)(gidr)	–1			–4

DE85 note that most words with final sonorant consonants as nuclei in fact optionally undergo prepausal annexation, yielding the variation illustrated in (35).

[7] As Karen Jesney (p.c.) points out, a serial analysis with separate nucleus projection and onset formation will require *CODA to overcome not only *C-NUC, but also ONSET, since the first step (rat)(lu)l(T) violates both. See further section 2.5.

(35) *Optional prepausal annexation*

(i)(gi)(dR) / (i)(gidr)   (R)(gL) / (Rgl)   (du)(mN) / (dumn)

By increasing the weight of *FINAL-NUC by 1, we get a tie in these cases.

(36) *Optional prepausal annexation as a tie*

/igidr/	*CODA 4	*FINAL-N 3	*C-NUC 1	
☞ (i)(gidr)	–1			–4
☞ (i)(gi)(dR)		–1	–1	–4

This is an approximation of an account of this instance of variation. Actual theories of variation in HG directly generate a probability distribution over candidates (“Max-Ent-OT”; Johnson 2002, Goldwater and Johnson 2003, Wilson 2006, Jaeger 2007), or use noise on the weights to generate variable outcomes across instances of evaluation (Boersma and Pater 2008). Both of these are in principle compatible with serial HG; the choice between them will likely depend on the success of associated learning algorithms, which are being developed in ongoing research.

In the current approximate account, we get the correct obligatory coda syllabification of a final obstruent, as shown in (37).

(37) *Obligatory final obstruent coda*

/ratlult/	*CODA 4	*FINAL-N 3	*C-NUC 1	
☞ (rat)(lult)	–1			–4
(rat)(lu)(lT)		–1	–3	–6

We also get the correct obligatory nuclear syllabification of a final vocoid (i, u, a), as shown in (38).

(38) *Obligatory final vocalic nucleus*

/tldi/	*CODA 4	*FINAL-N 3	*C-NUC 1	
(tLdj)	–1			–4
☞ (tL)(di)		–1		–3

The existence of gang effects in HG creates the possibility of using a single constraint to encode both the general preference for high sonority nuclei in Berber, as well as contribution of sonority to the explanation of the related absolute restriction against phrase-final obstruent nuclei. Again, no previous analysis has succeeded in formally drawing that connection.

2.3 Serial HG

Section 1.2 discussed the role of serialism in achieving a satisfactory account of ITB syllabification, and the last two sections explained the role of constraint weighting. We can now put these together in a serial HG derivation for /ratlult/. As Legendre, Sorace and Smolensky (2006) point out in their translation of the PS93 analyses into parallel HG, *HIATUS must have a value greater than the highest possible violation score on *C-NUC. Since we are assuming that this is maximum is 3, then as for *CODA, a weighting of 4 for *HIATUS will suffice.

The constraint weights are shown in (39), followed by the derivation. The optimum for each step of the derivation is shown in the left column, and the failed candidates are shown to its right. Underneath each candidate is its vector of violation scores in parentheses, followed by the weighted sum of violations. To save space, I have used positive integers for violation, and the weighted sum is also positive. The optimum hence has the lowest penalty. I have collapsed nucleus projection and onset adjunction into a single step, since this has no consequence here.[8]

(39)	*HIATUS	*CODA	*FINAL-NUC	*C-NUC
	4	4	3	1

Optima	*Failed candidates*			
1. (ra)tlult	r(aT)lult	ra(tL)ult	rat(lu)lt	ratl(wL)t
(0,0,0,0)=0	(0,0,0,3)=3	(0,0,0,1)=1	(0,0,0,0)=0	(0,0,0,1)=1
	ratlu(lT)			
	(0,0,1,3)=6			
2. (ra)t(lu)lt	(rat)lult			
(0,0,0,0)=0	(0,1,0,0)=4			
3. (rat)(lu)lt	(ra)t(lu)(lT)	(ra)t(lu)(L)t		
(0,1,0,0)=4	(0,0,1,3)=6	(1,0,0,1)=5		
4. (rat)(lul)t	(rat)(lu)(lT)			
(0,1,0,0)=4	(0,0,1,3)=6			
5. (rat)(lult)	(rat)(lul)(T)			
(0,0,0,0)=0	(0,0,0,3)=3			

[8] It turns out that the separation of core syllable construction into separate steps is desirable not only for reason of principle (see section 1.2), but may also be empirically necessary once we consider a fuller constraint set. If the constraint driving syllabification penalizes every unparsed segment (PS93's PARSE-SEG), and if core syllable construction is a single step, it will prefer the syllabification of a pair of segments over the syllabification of a single one (see relatedly Pruitt 2008 on foot parsing). In Berber, PARSE-SEG must have sufficient strength to overcome any degree of violation of *C-NUC, so as to rule out candidates that leave consonantal nuclei unparsed. As Kathryn Pruitt (p.c.) points out, this would result in (rŠ)q being chosen in step 1 of (22), leading to the ill-formed outcome of *(rŠq).

The first step is shown in the first row. Syllabifying either vowel as a nucleus violates none of our constraints, and is thus optimal. A full analysis of ITB would distinguish between high and low vowels, and would pick (ra)tlult over rat(lu)lt. The other failed candidates shown in step 1 include the full set of possible core syllables. All of the others have a consonantal nucleus, and thus a non-zero score as the last component of the violation vector, and a non-zero penalty. The final candidate ratlu(lT) has the worst nucleus in the worst (final) position, and thus has the highest penalty of (3✕1) + (1✕3) = 6.

In the second step we have an established syllable, and hence the option of adding a coda. This option is sub-optimal, however, relative to the creation of a second syllable, which again violates no constraint. In the third step, however, coda creation is optimal relative to the creation of a third new syllable, one that violates both *FINAL-NUC and has a stop nucleus. The fourth and fifth steps show the creation of the final coda. The optimum in the fourth step in fact ties with the optimum in the third, and the derivation could go in either order. In the fifth step I assume that the addition of the final consonant in the optimal candidate does not add a violation of *CODA ("every syllable ends in a vowel") and the creation of the final onsetless syllable does not violate *HIATUS ("no adjacent nuclei").

2.4 Serialism and HG typology

In the last few sections, the greater power of weighted than ranked constraints allowed *C-NUC to be used in the account of attested patterns that would fall out of reach of OT with the same constraint set. PS93, as well as Prince and Smolensky (1997) and Legendre, Sorace and Smolensky (2006), cite this greater power as the fatal flaw of a weighted constraint version of OT, claiming that it leads inexorably to unsupported typological predictions (though cf. Jesney 2009, Pater 2009, and Potts *et al.* 2010).

As pointed out in Pater (2009), a switch to serialism has deep consequences for HG/OT comparisons. Here I expand on that point with a further example and more discussion. The example relates to the process of coda syllabification that we have just seen in action in Berber. Pater, Bhatt and Potts (2007), drawing on Prince and Smolensky (1997), point out that a parallel version of HG can produce a system in which *CODA is satisfied at the cost of *n* violations of a faithfulness constraint, but not *n+1*. As Pater *et al*. (2007) point out, this example also crucially depends on a definition of *CODA that assigns a single violation for the entire coda, rather than one that assigns a violation per segment contained in the coda; the violation profiles in (40) reflect that definition.

To make the example similar to the one in the last section, we can use as our conflicting constraint PARSE-SEG, which assigns one violation per unparsed segment (PS93). A segment that violates PARSE-SEG may be unpronounced, as in PS93, or pronounced but unsyllabified; the choice is tangential to this example. As the pair of tableau in (40) shows, the result of the interaction of these constraints in parallel HG can be quite bizarre: a language that has codas with two or more consonants (e.g. ☞ [(apt)]), but not one (e.g. *[(ap)]). Further, this cut-off can be made at any point: languages with codas with no fewer consonants than three, or four, or five, or any other number can be modeled in this theory.

(40) *A problematic type of pattern in parallel HG*

ap	*Coda 1.5	Parse-Seg 1	
☞ (a)p		–1	–1
(ap)	–1		–1.5

apt	*Coda 1.5	Parse-Seg 1	
(a)pt		–2	–2
(ap)t	–1	–1	–2.5
☞ (apt)	–1		–1.5

This pattern is impossible in serial HG because of the limit to a single application of an operation in creating a candidate. The relevant operation here is adjunction, which adds a single consonant to an existing syllable. With our two constraints, the first application of adjunction will beat the fully faithful candidate if and only if *Coda has a greater weight than Parse-Seg. The tableau in (41) illustrates the second step of the derivation for the UR /apt/ with the same weights as in (40). Here we already have the nucleus syllabified through the prior application of nuclear projection. Importantly, the candidate set includes only the faithful candidate and the single adjunction, and not the fully syllabified candidate (apt) that was optimal in the parallel HG tableau for /apt/ in (40). Since that candidate is missing from the candidate set, the optimum is now [(a)pt]; we no longer get the strange pattern in which a coda is formed only to syllabify some minimum number of segments.

(41) *Coda formation in serial HG*

(a)pt	*Coda 1.5	Parse-Seg 1	
☞ (a)pt		–2	–2
(ap)t	–1	–1	–2.5

We can compare the typological predictions of serial and parallel versions of OT and HG using Staubs *et al.*'s (2010) OT-Help 2. The table in (42) provides an illustration of the ways in which the move to serialism can affect comparisons between ranked and weighted constraints. The top row displays a set of three inputs with potential codas of different lengths, and the subsequent rows show the syllabifications for four different languages. Checkmarks indicate which of the theories of constraint interaction can generate these languages with the two constraints *Coda and Parse-Seg. Serial candidate sets are produced using the adjunction operation as illustrated here, while the parallel candidate sets include codas with all possible numbers of consonants, from zero up to the number provided in the UR. Serial OT, parallel OT and serial HG each generate only two languages: one in which all of the consonants are parsed regardless of the number available in the UR, and one in which none of the consonants are. Parallel HG can also generate the extra languages alluded to above, in which a lower bound is placed on coda size: a coda is formed to parse minimally two, or three segments.

(42)

/ap/	/apt/	/aptk/	Parallel OT	Parallel HG	Serial OT	Serial HG
(a)p	(a)pt	(a)ptk	✓	✓	✓	✓
(ap)	(apt)	(aptk)	✓	✓	✓	✓
(a)p	(ap)t	(aptk)		✓		
(a)p	(a)pt	(aptk)		✓		

This example also illustrates a more general consequence of the single change limitation on candidates: that the set of possible constraint interactions, or "trade-offs" in the terminology of Pater (2009), is restricted relative to a parallel model. In particular, trade-offs involving multiple instances of violation of a given faithfulness constraint will never occur in a serial version of HG or OT, insofar as the operations that create candidates incur at most one violation of each faithfulness constraint (as in McCarthy 2007b). Thus, scenarios in which a markedness constraint is satisfied at the cost of *n* violations of a faithfulness constraint, but not *n+1*, which can be created in a parallel version of HG, are impossible in a serial version. The typological benefit accrued by serialism in the *CODA/PARSE example is likely typical of such cases.

2.5 Further issues

It is important to note that under certain assumptions serialism can also lead to unwelcome typological predictions, in both OT and HG. Based on results of a typological calculation in OT-Help, Pratt (2008) identifies the following derivational path as producing an outcome in serial OT that is impossible in parallel OT, a difference that seems to weigh in parallelism's favor. The last two steps of the derivation are show in (43). Prior to this, [a] was selected as the best nucleus, and [b] was added as its onset. In the third step, shown in (43a.), the coda [t] is chosen over the nuclear [n] because the single step of nucleus projection yields an onsetless syllable (ONSET is violated by a nucleus adjacent to a left syllable boundary). In the fourth step, the [n] is syllabified as a nucleus (assuming a ranking of *COMPLEX over *C-NUC).

(43) *Partial serial OT derivation*

a. Step 3

(ba)tn	ONSET	*CODA
☞ (bat)n		1
(ba)t(N)	1	

b. Step 4

(bat)n	ONSET	*CODA
☞ (bat)(N)		

The ONSET >> *CODA ranking thus yields onsetless syllables, which is quite counterintuitive, and almost certainly empirically problematic. In the full pattern that this derivational interaction produces, only relatively marked nuclei in the second syllable have no onsets; a UR like /tnba/ in which the last segment is a relatively good nucleus would be syllabified as (tN)(ba). Syllables that are marked on one dimension (nucleus sonority) are marked on another (onset possession).

Whether serial OT/HG produces other instances of such "markedness accumulation" is a question for further research that may bear on the choice between serialism and parallelism.

One solution to this particular problem is to introduce an operation of resyllabification that changes (bat)(N) into (ba)(tN). So long as this operation is cost-free, the candidate that it produces would always triumph, since it eliminates violations of *CODA and ONSET, and adds no other violations. The resyllabification operation is defined in (44).

(44) RESYLLABIFICATION
Input: A segment syllabified as a dependent in syllable 1, adjacent to a segment syllabified in syllable 2
Output: Segment adjoined to syllable 2

e.g. (bat)(N) → (ba)(tN) (ba)(tN) → (bat)(N)

To check that adding Resyllabification produces the desired result, I submitted to OT-Help the two inputs /bata/ and /batn/, with constraints PARSE-SEG, *C-NUC, *CODA, and *COMPLEX. In one typology calculation, I had only the Nucleus-Projection and Dependent Adjunction operations from (15), in the second I added Resyllabification. In both, I limited Adjunction to consonants. Serial OT and HG produced the same results, and so are collapsed in the table in (45).

(45)

/bata/	/batn/	Serial OT/HG no Resyllab.	Serial OT/HG w/ Resyllab.
bata	batn	✓	✓
(ba)(ta)	(ba)tn	✓	✓
(ba)(ta)	(bat)n	✓	✓
(ba)(ta)	(batn)	✓	✓
(ba)(ta)	(ba)(tN)	✓	✓
(ba)(ta)	(bat)(N)	✓	

This calculation shows that the addition of Resyllabification does produce the intended result: the theory can no longer produce the final language. The languages that are produced also seem typologically sensible: one language allows sonorant nuclei, as in [(ba)(tN)], while others avoid them by having instead complex codas, as in [(batn)], or by non-parsing, as in [(bat)n] or [(ba)tn]. The remaining potentially troubling case is the first language, which parses no segments at all, which again happens because the initial step of nucleus projection creates an onset violation (and ONSET overrides PARSE-SEG). This result could be avoided if contrary to one of this paper's basic premises, nuclear syllabification and onset formation were combined into a single step (*cf*. fn. 9 above), or if initial nucleus projection did not violate ONSET. However, a language in which the optima are unparsed is likely to be predicted by any version of OT/HG with a reasonably rich constraint set, if PARSE-SEG and other parsing constraints are ranked beneath all conflicting Markedness constraints.

Besides markedness accumulation, another potentially general problematic pattern of constraint interaction in Harmonic Serialism is one identified in Pater *et al*. (2007: 4.5), which can be termed “multiple application blocking”. When a markedness constraint can require multiple applications of an operation to be satisfied, we can get patterns in which the operation applies only in a situation in which it in fact satisfied by a single application. An example related to syllabification is the deletion of segments from coda position. As shown in (46), if *CODA assigns a violation to every syllable that ends in a consonant (PS93), deletion of a single segment will satisfy *CODA only if there is a single consonant in coda position.

(46) *A problematic type of pattern in Harmonic Serialism*

(ap)	*CODA	MAX
☞ (a)		1
(ap)	1	

(apt)	*CODA	MAX
(ap)	1	1
☞ (apt)	1	

Just like the example of markedness accumulation above, there is a variety of ways that this particular example of multiple application blocking might be ruled out. First, it may be that the inputs to these tableau could never occur: if *CODA is more highly valued than MAX, then each time a consonant is added to a syllable in coda position, it will be deleted, and we would never get to (apt). For this explanation to go through, it would also have to be the case that syllabification is universally absent from URs (see relatedly McCarthy and Pruitt to appear). As Pater *et al*. note, one might also simply redefine *CODA to penalize every segment in coda position, though a more satisfactory solution would rule this sort of interaction out in principle.

Unlike some other work in Harmonic Serialism that continues to assume parallelism for prosody (see *esp*. McCarthy 2007b), this paper assumes that both segmental and prosodic structure are constructed and revised serially (see also McCarthy 2008b, Pruit 2008, Elfner 2009, Jesney to appear, Kimper to appear). In the analysis of ITB, the absence of derivational look-ahead was crucial to achieving the correct outcome (see the discussion following (18) above). Another important question for further research is the extent to which serial prosodification can generally be maintained. It appears that none of the arguments against “bottom-up constructionism” in PS93 are in fact arguments against Harmonic Serialism (see e.g. McCarthy 2008b on Tongan). It also appears and that the candidate comparison that occurs at each step of the derivation gives the theory sufficient look-ahead to deal with these and many other cases that have been labeled as requiring parallelism (e.g. stress and syllabification in English; Pater 2000). There do, however, seem to be cases that are true problems for the most strictly serial account of prosodification; see McCarthy (2010) for an example and discussion (and Bakovic 2007 for some further relevant examples). Developing a general theory of operations that yields just the right amount of look-ahead is another important topic for further research.

3. Conclusions

In this paper I have proposed a grammatical framework that adopts serial candidate generation and evaluation with weighted constraints, serial HG. I have provided an analysis of a subset of the data on ITB syllabification in this framework, showing that the combination of serialism and constraint weighting allows for the first time an account of the ITB preference for high sonority nuclei that extends to related hard restrictions on nuclear sonority in Berber and elsewhere. Much remains to be done in terms of the development of this framework and this analysis, but these initial results seem promising for the future of serial HG, especially since serialism also imposes restrictions on the generative power of weighted constraint interaction.

References

Bakovic, Eric. 2007. A revised typology of opaque generalizations. *Phonology* 24, 217-259.

Staubs, Robert, Michael Becker, Christopher Potts, Patrick Pratt, Robert Staubs, John McCarthy and Joe Pater. 2010. *OT-Help 2: Java tools for Optimality Theory*. Software available at http://web.linguist.umass.edu/~OTHelp/

Blevins, J. 1995. The Syllable in Phonological Theory, *Handbook of phonological theory*, ed. by John Goldsmith, Basil Blackwell, London, 206-44.

Boersma, Paul and Pater, Joe. 2008. Convergence properties of a gradual learning algorithm for Harmonic Grammar. Ms, University of Amsterdam and UMass Amherst. Available at http://roa.rutgers.edu/.

Charette, Monique. 1991. *Conditions on phonological government*. Cambridge: Cambridge University Press.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In *Papers in Laboratory Phonology I*, ed. John Kingston and Mary Beckman, 283-333. Cambridge University Press.

Clements, G.N. 1997. Berber syllabification: derivations or constraints? In I.M. Roca, ed.. *Derivations and Constraints in Phonology*, Oxford: Oxford University Press, 289-330

Dell, Francois, and Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7:105–130.

Dell, Francois and Mohamed Elmedlaoui. 1988. Syllabic consonants in Berber: Some new evidence. *Journal of African Languages and Linguistics* 10, 1-17.

Dell, Francois and Mohamed Elmedlaoui. 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer Academic Publishers.

Elfner, Emily. 2009. Syllabification and Stress-Epenthesis Interactions in Harmonic Serialism. Ms, University of Massachusetts, Amherst.

Flack, Kathryn. 2009. Constraints on onsets and codas of words and phrases. *Phonology* 26: 269-302.

Flemming, Edward. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7–44.

Goldsmith, John. 1991. Phonology as an intelligent system. In *Bridges between Psychology and Linguistics: A Swarthmore festschrift for Lila Gleitman*, ed. Donna Jo Napoli and Judy Anne Kegl, 247–267. Hillsdale, NJ: Erlbaum.

Goldsmith, John. 1993a. Harmonic phonology. In The Last Phonological Rule, ed. John Goldsmith, 21–60. Chicago: University of Chicago Press.

Goldsmith, John. 1993b. The Last Phonological Rule: Reflections on Constraints and Derivations. Chicago: University of Chicago Press.

Goldsmith, John. 1994. A dynamic computational theory of accent systems. In Perspectives in phonology, ed. Jennifer Cole and Charles Kisseberth, 1–28. Stanford: CSLI.

Goldwater, Sharon and Johnson, Mark. 2003. Learning OT constraint rankings using a maximum entropy model. In Spenader, J., Eriksson, A., and Dahl, Ö., editors, Proceedings of the Stockholm Workshop on Variation within Optimality Theory, pages 111–120. Stockholm University, Stockholm.

Hankamer, Jorge and Judith Aissen. 1974. The sonority hierarchy. In A. Bruck, A. R. Fox and M. W. La Galy (eds.) Papers from the parasession on Natural Phonology. Chicago: Chicago Linguistic Society; 131-145.

Jäger, Gerhard. 2007. Maximum Entropy Models and Stochastic Optimality Theory, in Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling, and Chris Manning (eds.), *Architectures, Rules, and Preferences. Variations on Themes by Joan W. Bresnan*, CSLI Publications, Stanford, 467-479.

Jesney, Karen. 2009. Licensing in multiple contexts: an argument for Harmonic Grammar. Paper presented at the 45th Meeting of the Chicago Linguistic Society (CLS 45). Available at http://people.umass.edu/kjesney/papers.html.

Jesney, Karen. To appear. Positional Faithfulness, non-locality, and the Harmonic Serialism solution. In Suzi Lima, Kevin Mullin & Brian Smith (eds.), Proceedings of the 39th Meeting of the North East Linguistics Society (NELS 39). Amherst, MA: GLSA. Available at http://roa.rutgers.edu/.

Johnson, Mark. 2002. Optimality-theoretic lexical functional grammar. In Stevenson, S. and Merlo, P., editors, *The Lexical Basis of Syntactic Processing: Formal, Computational and Experimental Issues*, pages 59–73. John Benjamins.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Malden, MA: Blackwell.

Kimper, Wendell. To appear. Locality and Globality in Phonological Variation. In *Natural Language and Linguistic Theory*.

Lacy, Paul de. 2004. Markedness conflation in Optimality Theory. *Phonology* 21, 1-55.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Can connectionism contribute to syntax? Harmonic Grammar, with an application. In Proceedings of the 26th regional meeting of the Chicago Linguistic Society, ed. M. Ziolkowski, M. Noske, and K. Deaton, 237–252. Chicago: Chicago Linguistic Society.

McCarthy, John J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. 2003. OT constraints are categorical. *Phonology* 20: 75–138.

McCarthy, John J. 2007a. Restraint of analysis. In *Freedom of analysis*, ed. Sylvia Blaho, Patrick Bye, and Martin Krämer, 203–231. Berlin and New York: Mouton de Gruyter.

McCarthy, John J. 2007b. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.

McCarthy, John J. 2008a. The gradual path to cluster simplification. *Phonology* 25, 271-319.

McCarthy, John J. 2008b. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory*, 26(3), 499-546.

McCarthy, John J. 2010. Studying GEN. *Journal of the Phonetic Society of Japan* 13(2), 3-12.

McCarthy, John J. and Alan Prince. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology, in Mercè Gonzàlez, ed., Proceedings of the North East Linguistics Society 24, GLSA, Amherst, MA. Pp. 333–379. Available at http://roa.rutgers.edu/.

McCarthy, John J., and Alan Prince. 1999. Faithfulness and identity in Prosodic Morphology. In The prosody-morphology interface, ed. René Kager, Harry van der Hulst, and Wim Zonneveld, 218–309. Cambridge: Cambridge University Press.

McCarthy, John J. and Kathryn Pruitt (to appear) Sources of phonological structure. In Ralf Vogel and Hans Broekhuis (eds.), *Derivation and Filtering*. London: Equinox Publishing.

Pater, Joe. 2000. Nonuniformity in English stress: the role of ranked and lexically specific constraints. *Phonology* 17:2. 237-274.

Pater, Joe. 2009. Weighted constraints in generative linguistics. Cognitive Science 33, 999-1035. Available at http://roa.rutgers.edu/.

Pater, Joe, Rajesh Bhatt and Christopher Potts. 2007. Linguistic Optimization. Ms, University of Massachusetts, Amherst. Available at http://roa.rutgers.edu/.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt and Michael Becker. 2010. Harmonic Grammar with Linear Programming: From linear systems to linguistic typology. *Phonology* 27, 77-117.

Prince, Alan. 1983. Relating to the Grid. *Linguistic Inquiry* 14. 19-100.

Prince, Alan. 1997. Stringency and anti-Paninian hierarchies. Handout from LSA Institute, Cornell University. Available at http://ling.rutgers.edu/people/faculty/prince.html.

Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. RuCCS Technical Report 2, Rutgers University, Piscateway, NJ: Rutgers University Center for Cognitive Science. Revised version published 2004 by Blackwell. Page references to the 2004 version.

Pratt, Patrick. 2008. *Calculating typologies in Harmonic Serialism*. Honor’s thesis, University of Massachusetts, Amherst.

Prince, Alan, and Paul Smolensky. 1997. Optimality: From neural networks to universal grammar. *Science*, 275, 1604–1610.

Pruitt, Kathryn. 2008. Iterative foot optimization and locality in stress systems. Ms, University of Massachusetts, Amherst. Available at http://roa.rutgers.edu/

Wilson, Colin. 2003. Analyzing unbounded spreading with constraints: marks, targets and derivations. Unpublished ms, UCLA.

Wilson, Colin. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. Cognitive Science, 30, 945–982.

Zec, Draga. 1995. Coda constraints and conditions on syllable weight. *Phonology* 12, 85-129.