What Changes and What Stays the Same:

Is Harmonic Serialism with Positive Constraints still Optimality Theory?

Wendell Kimper || University of Manchester
wendell.kimper@manchester.ac.uk

27 May 2015|| mfm Fringe: W(h)ither OT?

1 Introduction

Spoiler alert: it's a trick question.

Optimality Theory (Prince and Smolensky, 1993/2004) gives us both a **theoretical framework** and a set of **specific theoretical tenets.**

It's not always possible to draw a sharp line between the two.

- As a framework, it's a platform for making predictions.
- As a theory, it advances claims that form the basis of those predictions.

My aim in this talk is to have a stab at untangling the framework and the theory.

2 The architecture of the framework, plus some specific theoretical tenets with proposed alternatives: a (non-exhaustive) whirlwind tour through recent history.

3 The interdependence of modifications to component parts.

4 Example: positive constraints in Harmonic Serialism

5 What are we trying to predict, anyway?

2 The Framework and the Theory

A barebones architecture for theories using the OT framework:

- Gen: A function mapping input representations to candidates
- Con: A set of violable constraints
- Eval: An optimisation function

Typographic conventions like small-caps constraints, tableaux, and manicules aren't part of either the framework or the theory, but they are useful and widely adopted.

Throughout this section, I'll discuss some specific theoretical claims advanced by Classic OT, with examples of alternative proposals. Neither is intended to be exhaustive, and the selection of examples is not meant to reflect either theoretical importance or empirical effectiveness.

2.1 Theories of Gen

Gen produces (/contains) all representationally possible forms (an infinite candidate set).[1]

- **cf.** OT-CC (McCarthy, 2007a) and Harmonic Serialism (McCarthy, 2000, 2007b), generation of contenders in Riggle (2004)

2.1.1 A Theory of Inputs

Richness of the Base: There are no language-specific restrictions on possible inputs.

- **cf.** ? (I'm not familiar with any work directly challenging this claim, but it probably exists.)

2.1.2 A Theory of Representations

Included with a candidate: correspondence relationships between input and output segments.

- **cf.** Containment theory in early OT (Prince and Smolensky, 1993/2004), coloured containment (van Oostendorp, 2007), Agreement by Correspondence (Hansson, 2001; Rose and Walker, 2004)

Classic OT doesn't advance any particular theory of features, feature geometry, etc.

2.2 Theories of Con

Universality: constraints in Con are universal (this may or may not require innateness).

- **cf.** Constraint induction (Albright and Hayes, 2002; Hayes and Wilson, 2008)

There are only Markedness constraints and Faithfulness constraints.

- **cf.** Targeted constraints (Wilson, 2004), procedural constraints (Blumenfeld, 2006), anti-faithfulness (Alderete, 2001), constraint conjunction (Smolensky, 1993, 1995)

Constraints assign *violation marks* (penalties) to candidates.

- **cf.** Positive constraints (Kimper, 2011)

2.2.1 A Theory of the Contents of Con?

Making proposals about which constraints are (or aren't) part of Con was the bread and butter of research in OT for a number of years, and continues to be part of the enterprise.

2.3 Theories of Eval

Constraints are ranked. In selecting a winner, Eval proceeds hierarchically, and domination is strict.

- **cf.** Cumulative constraint interaction in Harmonic Grammar (Smolensky and Legendre, 2006) and Maximum Entropy (Goldwater and Johnson, 2003)

[1] A koan: if Gen does everything, is Gen doing anything?

Eval returns a single optimal candidate.

- **cf.** Coetzee's (2006) well-formedness ordering, MaxEnt's probability distribution

The optimal candidate selected by Eval exits the derivation as the surface (phonological) form.

- **cf.** Harmonic Serialism, Stratal OT (Kiparsky, 2000)

3 Interdependence

Some types of interdependence among components is obvious: the restricted Gen of Harmonic Serialism requires an Eval that does not always feed to the surface.

Others are not quite as obvious, but non-trivial.

3.1 Implications for The Contents of Con

With identical constraint sets, the grammars produced by HG are a superset of those produced by OT.

- But this is not necessarily true if the contents of Con are also subject to modification.

Jesney (2011): in OT, both positional licensing and positional faithfulness are needed, but in HG it may be possible to get the same empirical coverage with only positional licensing.

In HS, restricting Gen to a single instance of a single change requires that we develop a clear theory of what our operations are, and the representations on which those operations are performed.

4 The Case of Positive Constraints

The goal of this section is to demonstrate the interdependence of OT's component parts with a concrete example: including positively-defined constraints in Con.

The question: do phonological processes happen in order to avoid problems, or can they also promote a desirable state of affairs?

- In the case of harmony: is the goal to avoid disharmonic sequences/segments, or to realise the spreading feature across multiple segments?

This has consequences for our predicted typologies.

- A negative harmony constraint predicts that languages should avoid disharmony, even when that doesn't result in more spreading.
- A positive constraint predicts languages should promote or preserve spreading, even when that doesn't reduce disharmony.

4.1 Gen and the Infinite Goodness Problem

Positive and negative constraints are not simply the mirror image of each other. Compare, for example, positive and negative formulations of a constraint like Onset:

(1) a. (−)Onset: Assign -1 for every syllable that does not have an onset.
b. (+)Onset: Assign +1 for every syllable that has an onset.

Both versions of the constraint demand that syllables have onsets; both prefer a candidate with an epenthetic onset (2a) over a faithful, onsetless candidate (2b).

- However, a positively formulated Onset constraint will prefer a candidate with both an epenthetic onset and an epenthetic CV syllable (2c) over a candidate with a single epenthetic onset (2b), while the negative version of the constraint exerts no such preference.

(2)

ap	(−)Onset	(+)Onset
a. .ap.	−1	
b. .ʔap.		+1
c. .ʔap.ʔə.		+2

The problem with this is immediately apparent: if (+)Onset favors a candidate with an additional epenthetic syllable over a candidate with onset epenthesis, it will favor a candidate with two epenthetic syllables over a candidate with only one epenthetic syllable, and so on.

- “There is no candidate that has the maximal value . . . and were the constraint asked to produce the class of forms that do maximally well on it, no output would be defined. If such a constraint is admitted, the theory ceases to exist.” (Prince, 2007)

(3)

ap	(−)Onset	(+)Onset
a. .ap.	−1	
b. .ʔap.		+1
c. .ʔap.ta.		+2
d. .ʔap.ta.ka.		+3
e. .ʔap.+$\aleph_0$		+∞

If the theory of Con is modified to allow constraints to assign rewards, while holding the theory of Gen constant, the situation is untenable.

If the theory of Gen is modified to produce a restricted candidate set, the infinite goodness problem can be avoided. Compare, the candidates generated as a result of epenthesis by Classic OT’s unrestrained Gen and those created by HS’s restricted Gen.

(4) *Possible candidates in OT and HS*

	OT	**HS**
Input	/ap/	.ap.
a.	.ap.	.ap.
b.	.ʔap.	.ʔap.
c.	.ʔap.ta.	—
d.	.ʔap.ta.ka.	—
e.	.ʔap.+$\aleph_0$	—

At any given step of the derivation, the candidate set produced by HS's restrained Gen is finite — there will be single optimum.

- It is still possible to define a positive constraint that provokes infinite goodness: a constraint may be viably positive `iff` the structure which earns the reward depends critically on some prerequisite condition whose creation requires a distinct derivational step.

4.2 Predictions

4.2.1 Predictions of a Negative Constraint

Wilson (2004, 2006) identifies a number of pathological predictions made by common harmony constraints. In particular, commonly-used constraints make problematic predictions in systems where spreading is blocked. For example, in Malay, nasal harmony is blocked by obstruents (Onn, 1980; Walker, 2000):

(5) *Nasal harmony in Malay*

a. Unbounded rightward spreading

mĩnõm	'to drink'
bãŋõn	'to rise'
mãj̃ãn	'stalk (palm)'
pəŋə̃ŋãh̃ãn	'central focus'
mə̃nãw̃ãn	'to capture'

b. Blocked by obstruents

mãkan	'to eat'
pəŋãw̃ãsan	'supervision'

Agree, which assigns violations to a sequence of adjacent segments with different values for a given feature (Bakovic, 2000; Eisner, 1999; Lombardi, 1999, 2001; Pulleyblank, 2004) suffers from the *Sour Grapes* problem — if complete harmony is impossible, no spreading is predicited.

Constraints like Align assign violations to segments that intervene between the edge of the feature domain and the edge of the bounding domain (Archangeli and Pulleyblank, 2002; Cole and Kisseberth, 1994, 1995; Kirchner, 1993; Pulleyblank, 1996; Smolensky, 1993).

A sufficiently highly ranked Align will block epenthesis. Consider Malay′:

(6)

nawakast	Align-R(nas,PWd)	*CC#	Dep
a. ☞ nãw̃ãkast	–4	–1	
b. nãw̃ãkasət	W–5	L	W–1

In Malay′, final clusters are normally resolved by epenthesis (7a). However, in words where spreading is blocked, no epenthesis occurs (7b).

(7)

	Align-R(nas,PWd)	*CC#	Dep
a. tawakasət ∼ tawakast		W	L
b. nãw̃ãkast ∼ nãw̃ãkasət	W	L	W

This is unattested. Wilson (2004) identifies a number of other similarly unattested predictions that constraints like align make (including those related to deletion, allomorph selection, and reduplication).

Broadly speaking, these predictions arise from the fact that negatively-defined harmony constraints are sensitive to *the number of non-harmonised segments* in a word.

- (...and this is still true in HS.)

4.2.2 Predictions of a Positive Constraint

A positive spreading like constraint (8) assigns rewards to harmonised segments rather than assigning penalties to non-harmonised ones.

(8) (+)Spread(F): Assign +1 for each segment linked to F as a dependent.

In HS, a derivation with this constraint proceeds as in (9).

- At the first step, the candidate with nasal spreading wins over the faithful candidate. This is again the case at the second and third steps; spreading will continue to iterate, segment by segment, until there are no more segments left unassociated or until spreading is blocked.

(9) **Step 1**

nawakast	*NasObs	(+)Spread(nas)	Ident(nas)
a. nawakast		W	L
b. ☞ nãwakast		+1	–1

Step 2

nãwakast	(*NasObs	(+)Spread(nas)	Ident(nas)
a. nãwakast		W+1	L
b. ☞ nãw̃akast		+2	–1

Step 3

nãw̃akast	(*NasObs	(+)Spread(nas)	Ident(nas)
a. nãw̃akast		W+2	L
b. ☞ nãw̃ãkast		+3	–1

Step 4: *Convergence*

nãw̃ãkast	*NasObs	(+)Spread(nas)	Ident(nas)
a. ☞ nãw̃ãkast		+3	
b. nãw̃ãk̃ast	W-1	L+4	W–1

(+)Spread(nas) will not be sensitive to the number of segments to the right of the blocking obstruent.

(10) **Step 4**

nãw̃ãkast	(+)Spread(nas)	*CC
a. nãw̃ãkast	+3	W-1
b. ☞ nãw̃ãkasət	+3	

A constraint like (+)Spread(nas) is prevented from *inducing* epenthesis by the restricted Gen in HS — epenthesis and harmony cannot occur concurrently.

(11) **Step 4**: *Convergence*

nãw̃ãkast	(+)Spread(nas)	Dep
a. ☞ nãw̃ãkast	+3	
b. nãw̃ãəkast	+3	W–1

A positive spreading constraint exerts no preference against non-harmonized segments, and because a HS Gen is restricted to making gradual changes, a positive constraint cannot itself induce processes that increase the number of segments available to harmony.

(+)Spread(nas) is prevented from inducing epenthesis, but it can still exert influence in processes that affect the number of assimilated segments.

- If Spread(F) outranks a constraint that produces vowel reduction, the predicted result is a language where reduction is blocked just in case the targeted segment is a dependent of a harmony domain.

A pattern like this is in fact attested in Kera (Pearce, 2008), where (front/back) harmony blocks a prosodically-driven process of reduction along the F_2 dimension.

(12) **Vowel Harmony in Kera**

a. Back/round harmony

/ciːr-i/	[ciːri]	'your (f) head'
/ciːr-u/	[cuːru]	'his head'

Kera also exhibits a process of vowel reduction — unstressed vowels shorten and reduce.

However, Pearce (2008, 2011) reports that vowels in a back/round harmony domain reduce in duration and along the F1 dimension, but *do not reduce along the F2 dimension.*

- cf. vowels which are not members of harmony domains, which reduce along all dimensions.

(13) **Step 1**

[+][−] cɨr-i	Spread(−back)	Reduce	Ident(back)
a. [+][−] cɨr-i	W	−1	L
b. [+] cɨr-ə	W	L	−1
c. ☞ [−] cir-i	+1	−1	−1

Step 2: *Convergence*

[−] cir-i	Spread(−back)	Reduce	Ident(back)
a. ☞ [−] cir-i	+1	−1	
b. [−] cir-ə	W	L	−1

The pattern of harmony seen in Kera is consistent with the predictions of a positive harmony constraint — members of a harmony domain are protected from a process that threatens to shrink that domain.

4.2.3 Summary

Harmony constraints that assign rewards make different predictions from those which assign penalties.

The predictions of positive harmony constraints seem to be a better fit for the range of attested harmony patterns than the predictions of negative harmony constraints.

But! Including positively defined constraints in a theory of Con requires also adopting a restricted theory of Gen in order to be viable — the two components can't be examined in isolation.

4.3 So... is this still OT?

That which we call a rose by any other name would smell as sweet.

5 The Problems of Prediction

The difficult question: What are we trying to make predictions about? And how do we measure success?

The universality of Classic OT's Con is designed to make predictions about typology.

- The Too Many Solutions Problem has often been cast as an issue with OT, but over- and under-generation are just mismatch between our predictions and (what we know of) reality.
- Rather than being the problem, OT's machinery allowed us to make falsifiable predictions. And some of them turned out to be false. This is a feature, not a bug.

However, there is increasing awareness in the field that:

- The set of known languages is a very poor statistical sample.
- Much of the data we've based our generalisations on is of dubious quality.

OT's usefulness in generating typological predictions has not changed, but we're finally realising that our metric for success in evaluating those predictions has some serious issues.

I still think generating typological predictions has some value.

- We don't have all the information we need to test them, but we still have *some* information.
- ...and the quality and extent of our primary phonological data is improving.

We can (and should) also use the framework of OT to generate predictions about things we are better able to test, such as behavioural data.

- See e.g. Daland et al. (2011) on sonority projection, Moreton et al. (2013) on learning.

We should also keep losing sleep over this question.

6 The Future?

It is very difficult to prophesy, especially about the future.

We have always treated the specific theoretical claims of OT as subject to falsification and revision.

But OT has left us with a broad framework for generating predictions that is highly flexible and increasingly computationally implementable.

- Why would we abandon that?

References

Albright, Adam, and Bruce Hayes. 2002. Modelling english past tense intuitions with minimal generalization. In *Proceedings of the Sixth Meeting of the ACL Special Interest Group in Computational Phonology*, ed. M. Maxwell. Association for Computational Linguistics.

Alderete, John. 2001. Dominance effects as trans-derivational anti-faithfulness. *Phonology* 18:201–253.

Archangeli, Diana, and Douglas Pulleyblank. 2002. Kinande vowel harmony: Domains, grounded conditions, and one-sided alignment. *Phonology* 19:139–188. University of Arizona and University of British Columbia.

Bakovic, Eric. 2000. Harmony, dominance, and control. Doctoral Dissertation, University of California, Santa Cruz. ROA-360.

Blumenfeld, Lev. 2006. Constraints on phonological interactions. Doctoral Dissertation, Stanford University.

Coetzee, Andries. 2006. Variation as accessing "non-optimal" candidates. *Phonology* 23:337–385.

Cole, Jennifer, and Charles W. Kisseberth. 1994. An optimal domains theory of harmony. Urbana: University of Illinois. ROA-22.

Cole, Jennifer, and Charles W. Kisseberth. 1995. Nasal harmony in Optimal Domains Theory. Ms., University of Illinois. ROA-49.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andreas Davis, and Ingrid Normann. 2011. Explaining sonority projection effects. *Phonology* 28:197–234.

Eisner, Jason. 1999. Doing OT in a straightjacket. Colloquium, University of California, Los Angeles.

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Workshop on Variation within Optimality Theory*, ed. Jennifer Spenader, Anders Eriksson, and Östen Dahl, 111–120. Stockholm University.

Hansson, Gunnar. 2001. Theoretical and typological issues in consonant harmony. Doctoral Dissertation, University of California, Berkeley.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic In* 39:379–440.

Jesney, Karen. 2011. Licensing in multiple contexts: an argument for Harmonic Grammar. In *Proceedings of the 45th Meeting of the Chicago Linguistics Soceity*.

Kimper, Wendell. 2011. Competing triggers: Transparency and opacity in vowel harmony. Doctoral Dissertation, University of Massachusetts Amherst.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17:351–367.

Kirchner, Robert. 1993. Turkish vowel harmony and disharmony: An Optimality Theoretic account. Ms., UCLA. ROA-4.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17:267–302.

Lombardi, Linda, ed. 2001. *Segmental Phonology in Optimality Theory. Constraints and Representations*. Cambridge University Press.

McCarthy, John. 2000. Harmonic Serialism and parallelism. In *Proceedings of NELS 30*, ed. Masako Hirotani, 501–24. Amherst, MA: GLSA Publications.

McCarthy, John. 2007a. *Hidden generalizations: Phonological opacity in optimality theory*. London: Equinox Publishing.

McCarthy, John. 2007b. Restraint of analysis. In *Freedom of analysis*, ed. Sylvia Blaho, Patrick Bye, and Martin Krämer, 203–31. Berlin and New York: Mouton de Gruyter.

Moreton, Elliott, Joe Pater, and Katya Pertsova. 2013. Phonological concept learning. Ms, University of North Carolina, Chapel Hill and University of Massachusetts, Amherst.

Onn, Farid. 1980. Aspects of Malay phonology and morphology. Doctoral Dissertation, Universiti Kebangsaan Malaysia.

van Oostendorp, Marc. 2007. Derived environment effects and consistency of exponence. In *Freedom of analysis?*, ed. P. Blye S. Blaho and M. Krämer. Berlin/New York: Mouton de Gruyter.

Pearce, Mary. 2008. Vowel harmony domains and vowel undershoot. In *Ucl working papers in linguistics*, volume 20. Research Department of Linguistics, University College London.

Pearce, Mary. 2011. Effects of atr, fronting and rounding harmony on reduction. Presented at the 19th Manchester Phonology Meeting.

Prince, Alan. 2007. The pursuit of theory. In *Cambridge handbook of phonology*, ed. Paul de Lacy. Cambridge University Press.

Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality theory: Constraint interaction in generative phonology*. Blackwell Press.

Pulleyblank, Douglas. 1996. Neutral vowels in Optimality Theory: A comparison of Yoruba and Wolof. *Canadian Journal of Linguistics* 41:295–347.

Pulleyblank, Douglas. 2004. Harmony drivers: no disagreement allowed. In *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*.

Riggle, Jason. 2004. Generation, recognition, and learning in finite state optimality theory. Doctoral Dissertation, University of California, Los Angeles.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80:475–531.

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Handout to Presentation at the Rutgers Optimality Workshop 1.

Smolensky, Paul. 1995. On the internal structure of the constraint component *Con* of ug. Handout to Presentation given at UCLA, April 1995.

Smolensky, Paul, and Géraldine Legendre. 2006. *The harmonic mind*. MIT Press.

Walker, Rachel. 2000. Long-distance consonant identity effects. In *Proceedings of the 19th West Coast Conference on Formal Linguistics*, ed. Roger Billerey and Brook Lillehaugen.

Wilson, Colin. 2004. Analyzing unbounded spreading with constraints: marks, targets, and derivations. UCLA.

Wilson, Colin. 2006. Unbounded spreading is myopic. Bloomington, IN: PhonologyFest 2006.