Why (not) weight? The case for Harmonic Grammar

Joe Pater, University of Massachusetts, Amherst

LSA Summer Institute, University of Colorado, Boulder
July 2011

1. Background

Weighted constraint models of grammar have received considerable recent attention because they can straightforwardly encode grades of well-formedness:

(1) Brown, 2008; Coetzee & Pater, 2008b; Hayes & Wilson, 2008; Keller, 2006; Martin, 2007; McClelland & Vander Wyck, 2006

Harmonic Grammar (HG) was in fact originally motivated as an account of gradient syntactic well-formedness (Legendre, Miyata, and Smolensky 1990; although cf. Legendre, Sorace and Keller 2006).

Weighted constraints are also attractive because they lend themselves well to producing variable outcomes:


And in both of these cases, a large part of the attraction is the availability of learning algorithms that cope with these two kinds of non-categoricity (cf. Tesar and Smolensky 2000).

In developing Optimality Theory (OT), Prince and Smolensky (1991) in fact originally pursued a version of their theory with weighted constraints. Why did they switch to ranked constraints? Prince and Smolensky (1993/2004: 232) cite a hypothetical interlocutor as having the following “fear of optimization”:

(3) Loss-of-restrictiveness: “In order to handle optimality, you must use numbers and use counting... The result will be a system of complicated trade-offs... giving tremendous descriptive flexibility and no hope of principled explanation. Therefore, the main goal of generative grammatical investigation is irredeemably undermined.”

Their response (p. 233) is that this “[c]oncern is well-founded,” but that in OT “recourse to the full-blown power of numerical optimization is not required....” Smolensky (2006a) states that “[c]onstraint interaction in human grammars is more restricted than that permitted by arbitrary numerical constraint strengths. The more restricted theory of constraint interaction that yields empirically valid typologies is [OT]....”
These claims about the relative restrictiveness of HG and OT are accompanied by little explicit comparison of differences between the typological predictions of weighted and ranked constraints. Prince and Smolensky (1993/2004) present no examples of unattested linguistic patterns generated just by constraint weighting, and in fact, there appear to be no such cases in the published literature prior to Legendre et al. (2006), who provide an example discussed in section 7 below.

There has recently emerged a body of research addressing the question of HG/OT differences, which has been greatly aided by the parallel emergence of computational tools to aid this work. The aim here is to provide a general guide to the situations in which HG and OT do, and do not, yield differences in their typological predictions.

2. The necessity of asymmetric trade-offs

We start with a simple HG tableau. The optimum is the candidate with the highest *Harmony*, which in HG is the weighted sum of constraint violations. (4) From Pater (2009a)

(9) Final devoicing in HG

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>/bad/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bad</td>
<td>−1</td>
<td>−2</td>
</tr>
<tr>
<td>pad</td>
<td>−1</td>
<td>−3</td>
</tr>
<tr>
<td><em>Coda-Voice</em> [bat]</td>
<td>−1</td>
<td>−1</td>
</tr>
<tr>
<td>pat</td>
<td>−2</td>
<td>−2</td>
</tr>
</tbody>
</table>

Demo: bad-excel.txt

The weights in the above tableau are partially arbitrary; there is an infinite set of weights that could be used to make [bat] optimal.

Discussion point: Why?

The non-arbitrary aspect is that for [bat] to be optimal, the weight of *Coda-Voice must be greater than that of Ident-Voice. With the reverse relationship, [bad] becomes optimal. Here the numerical weights are behaving just like OT’s ranked constraints.

Given a desired optimum, or “Winner”, a set of failed candidates, or “Losers”, and their associated vectors of scores on a set of constraints, we can thus produce a set of linear inequalities, or weighting conditions, that must obtain if the Winner is to be made correctly optimal (cf. Prince’s 2002 OT ranking conditions). We can obtain a useful representation for examining weighting conditions by subtracting the scores
of the Winner from those of a Loser, thus producing an HG comparative vector (cf. again Prince 2002 in OT). The comparative vectors from the candidates in (4), with [bat] as the Winner, are shown in (5).

(5) **Comparative vectors for (4)**

<table>
<thead>
<tr>
<th>W ~ L</th>
<th>*Coda-Voice</th>
<th>Ident-Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>[bat] ~ [bad]</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>[bat] ~ [pat]</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>[bat] ~ [pad]</td>
<td>+1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Weighting conditions from comparative vectors:

(6) a. For the Winner to be correctly optimal, the sum of the scores in each row, each times the constraint’s weight, must be greater than zero.

b. When all of the non-zero scores are +1 and −1, as in (5), we can also simply say that the sum of the weights of the constraints preferring the Winner must be greater than the sum of the weights preferring the Loser.

Potts et al. (2009) show that Linear Programming’s simplex algorithm can be used to find constraint weights that meet these weighting conditions, and also to detect when no such weighting exists. As such, it does for HG what Recursive Constraint Demotion (RCD, Tesar and Smolensky 1998) does for OT. OT-Help (Staubs et al. 2010) uses Potts et al.’s LP method, as well as RCD, to find correct HG

**Discussion point:** Harmonic bounding and restriction of weights to positive values.

In this example, violations trade one-to-one (see Prince 2002 on one-to-one trade-offs in HG–OT translations):

(7) Each potential voiced coda incurs either one violation of *Coda-Voice if it is voiced in the output, or one violation of Ident-Voice if it is voiceless.

Thus, for every violation of *Coda-Voice that a candidate avoids, it will incur one violation of Ident-Voice, and vice versa, as illustrated in (8).

(8) **Symmetric trade-off**

<table>
<thead>
<tr>
<th>/dagbad/</th>
<th>*Coda-Voice</th>
<th>Ident-Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>dag.bad</td>
<td>−2</td>
<td></td>
</tr>
<tr>
<td>dag.bat</td>
<td>−1</td>
<td>−1</td>
</tr>
<tr>
<td>dak.bat</td>
<td></td>
<td>−2</td>
</tr>
</tbody>
</table>

In other words, violations trade off symmetrically. To get a difference between HG and OT, we must have an asymmetric trade-off, as in the following example from
Japanese loanword phonology (Kawahara, 2006; Nishimura, 2003, 2006; all data are from the former source).

(9) **Violations of Lyman’s Law in loanwords**

[bagi:] ‘buggy’  [bagu] ‘bug’  
[bogi:] ‘bogey’  [dagu] ‘Doug’  
[bobu] ‘Bob’  [giga] ‘giga’

(10) **Voiced/voiceless obstruent near-minimal pairs**

[sunob:u] ‘snob’  [sutop:u] ‘stop’  
[kid:o] ‘kid’  [kit:o] ‘kit’  
[red:o] ‘red’  [autoret:o] ‘outlet’  
[hed:o] ‘head’  [met:o] ‘helmet’

(11) **Optional devoicing of a geminate in the Lyman’s law environment**

[gud:o] ~ [gut:o] ‘good’  
[bed:o] ~ [bet:o] ‘bed’  
[dored:o] ~ [doret:o] ‘dredlocks’  
[bad:o] ~ [bat:o] ‘bad’  
[deibid:o] ~ [deibit:o] ‘David’

[dog:u] ~ [dok:u] ‘dog’  
[bag:u] ~ [bak:u] ‘bag’  
[bud:a] ~ [but:a] ‘Buddha’  
[dorag:u] ~ [dorak:u] ‘drug’  
[big:u] ~ [bik:u] ‘big’

Demo: [japanese-OT-Help.txt](#) in Excel and OT-Help

**Discussion points:** OT-Help file format, OT-HG typological differences, comparative views, ranking and weighting conditions.

Demo: [japanese-max-ent-excel.xls](#)

**Discussion point:** Maximum Entropy Grammar: a probabilistic variant of HG (e.g. Goldwater and Johnson 2003, Hayes et al. 2009).

We can also get an asymmetric trade-off from just two constraints, if one of those constraints assigns violations gradiently, in McCarthy’s (2003) sense. McCarthy suggests the following restatement of Prince and Smolensky’s (1993/2004) H-Nuc, which I call *C-Nuc in Pater (to appear)*

(12) *C-Nuc
Assign a violation mark to a nucleus for each degree of sonority separating it from [a]

4
McCarthy rejects this constraint, and all scalar constraints, as OT constraints because they cannot produce attested patterns (see also P&S ch. 9). When \( \ast \text{-Nuc} \) conflicts with another constraint, OT can produce only two outcomes:

\[
\begin{align*}
\ast \text{-Nuc} >> \ast \text{Complex} & \quad \text{A cluster is better than any consonantal nucleus} \\
\ast \text{Complex} >> \ast \text{-Nuc} & \quad \text{Any consonantal nucleus is better than a cluster}
\end{align*}
\]

\[
\begin{align*}
\ast \text{-Nuc} >> \text{Dep} & \quad \text{No consonantal nuclei permitted (epenthesis everywhere)} \\
\text{Dep} >> \ast \text{-Nuc} & \quad \text{All consonantal nuclei permitted (epenthesis nowhere)}
\end{align*}
\]

In HG, however, we can get intermediate outcomes (nasal gets one violation, fricative 2):

\[
\begin{array}{c|cc|c}
\text{Dep} & \ast \text{-Nuc} & \text{tN} & \text{tVn} \\
1.5 & 1 & -1 & -1.5 \\
\hline
\text{ts} & \text{Dep} & \ast \text{-Nuc} & \text{tS} & \text{tVs} \\
1.5 & 1 & -2 & -1 & -1.5
\end{array}
\]

**Discussion points:** Weighting conditions, HG typology.

An asymmetric trade-off is a precondition for a *gang effect*, or *cumulative constraint interaction* (Pater 2099: 1002):

\[
\begin{align*}
\text{(15)} & \quad \text{A constraint is satisfied at the cost of } n \text{ violations of some lower valued constraint(s), but not } n + 1.
\end{align*}
\]

An intuitive way of thinking about asymmetric trade-offs and gang effects ("One" is really \( n \) and "Two" is really \( n + 1 \)):

\[
\begin{align*}
\text{(16)} & \quad \text{To get a gang effect, it must be the case that a violation of one constraint can be used to avoid two violations of some other constraint(s).}
\end{align*}
\]

When we can’t have gang effects, we can’t get HG-OT differences. While this hopefully seems straightforward, a failure to fully appreciate the need for asymmetric trade-offs is often at the root of imagined consequences of adopting weighted constraints in OT.

In making the case for OT to an interdisciplinary audience, Prince and Smolensky (1997: 1604) draw the generalization that:

\[
\begin{align*}
\text{(17)} & \quad \text{In a variety of clear cases where there is a strength asymmetry between two conflicting constraints, no amount of success on the weaker constraint can compensate for failure on the stronger one.}
\end{align*}
\]
They attribute this type of phenomenon to strict domination property of ranked constraints. As an example, they discuss the interaction of NoCoda and Parse (Parse is a faithfulness constraint that demands that input segments be parsed into output syllable structure; a consonant that is unparsed is unpronounced (=deleted)). Prince and Smolensky (1997: 1606) state that:

(18) Domination is clearly “strict” in these examples: No matter how many consonant clusters appear in an input, and no matter how many consonants appear in any cluster, [the grammar with NoCoda >> Parse] ... will demand that they all be simplified by deletion (violating Parse as much as is required to eliminate the occasion for syllable codas), and [the grammar with Parse >> NoCoda] ... will demand that they all be syllabified (violating NoCoda as much as is necessary). No amount of failure on the violated constraints is rejected as excessive, as long as failure serves the cause of obtaining success on the dominating constraint.

In this passage, Prince and Smolensky offer as an illustration of strict domination the observation that the number of consonant clusters (i.e., potential codas) does not affect whether deletion occurs or not. The table in (19) shows that the trade-offs across candidates' violation profiles have the same symmetry as the coda devoicing case in (8). Unparsed segments are placed between angled brackets; as they are unparsed, they do not violate NoCoda.

(19) Another symmetric trade-off

<table>
<thead>
<tr>
<th>/dagbadga/</th>
<th>NoCoda</th>
<th>Parse</th>
</tr>
</thead>
<tbody>
<tr>
<td>dag.bad.ga</td>
<td>−2</td>
<td></td>
</tr>
<tr>
<td>da&lt;g&gt;.bad.ga</td>
<td>−1</td>
<td>−1</td>
</tr>
<tr>
<td>da&lt;g&gt;.ba&lt;d&gt;.ga</td>
<td>−2</td>
<td></td>
</tr>
</tbody>
</table>

Discussion point: What about the number of consonants in a single cluster?

Demo: CVCCCC-1.txt and CVCCCC-2.txt

3. The insufficiency of asymmetric trade-offs

While an asymmetric trade-off is a necessary condition for an HG-OT difference, it is not a sufficient one.

Demo: bad-OT-help.txt

*Voice and *Coda-Voice can act in a gang effect in choosing [bat] for /bad/, but that gang effect is vacuous in that it does not lead to an HG-OT difference in typological predictions.

Pace Pater (2009a), it is not the case that gang effects between constraints in a specific-to-general or “stringency” (Prince 1997) relationship will always be vacuous. The universal vacuity claim is an overgeneralization because even though the general constraint will always assign a violation whenever the specific one does, the competing candidate might itself violate the general constraint, leading to a difference between the candidates in only the violation of the specific constraint. This possibility is realized in the following scenario.

The constraints in the specific-to-general relation are *Stress-[i], which assigns violations to stressed high vowels, and *Stress-[i,e], which assigns violations to stressed high and mid vowels (de Lacy 2004). The competing constraint Stress-Final demands that stress be on the final syllable. The first tableau contains the pair of candidates whose violations on the specific and general pair differ only on the specific *Stress-[i]; the shared violation of general *Stress-[i,e] is canceled out.

(20) A non-vacuous specific-to-general gang effect

<table>
<thead>
<tr>
<th>/teni/</th>
<th>Stress-Final</th>
<th>*Stress-[i]</th>
<th>*Stress-[i,e]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ténɪ/</td>
<td>1.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>[ténɪ]</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>☞ [tenɪ]</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/tane/</th>
<th>Stress-Final</th>
<th>*Stress-[i]</th>
<th>*Stress-[i,e]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/táne/</td>
<td>1.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>[táne]</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>☞ [tané]</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/tani/</th>
<th>Stress-Final</th>
<th>*Stress-[i]</th>
<th>*Stress-[i,e]</th>
</tr>
</thead>
<tbody>
<tr>
<td>☞ [táni]</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>[taní]</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

The gang effect is seen in the final tableau, in which violation of both lower valued constraints is worse than a violation of the higher valued one. This gang effect is non-vacuous in that the overall pattern is different from the one that would be produced if only the specific *Stress-[i] constraint were active in selecting optima, as it would be if its weight/rank were higher than Stress-Final, and the weight/rank of *Stress-[i,e] were lower than Stress-Final. In that case, stress would be uniformly initial when the final syllable contained [i], and the initial syllable contained [a] or [e]. In the pattern illustrated in (2), stress fails to retract to [e] because *Stress-[i] is weighted beneath Stress-Final (see the first tableau). Only when the candidates
differ on both *Stress-[i] and *Stress-[i,e], as in the last tableau, will stress leave its preferred final position.

Thus, under the right conditions, constraints in a specific-to-general relation can produce the "sufficient reward" threshold that distinguishes weighted from ranked constraint interaction: a general preference is overridden only to gain a sufficient benefit on another dimension. Here, that benefit is stressing the best type of vowel (low [a]) instead of the worst one (high [i]). A gain from worst to intermediate (high [i] to mid [e]) or intermediate to best (mid [e] to low [a]) is insufficient to compensate for placing stress on the dispreferred non-final position. De Lacy’s (2004, 2006) typological survey appears to include no vowel quality-based stress pattern with a sufficient reward threshold. Further research is required to determine whether that gap is accidental, as predicted by HG, or is a reflection of a general restriction on constraint interaction, as predicted by OT.

As another type of example of a vacuous gang effect, we turn to a case of constraint interaction that Prince and Smolensky (1993/2004) point to as illustrating the difference between ranking and weighting. It involves the pair of candidates in their tableau (183A), which forms part of their analysis of Lardil final vowel truncation (Hale, 1973):

\[(21)\] From Pater (2009a)

<table>
<thead>
<tr>
<th>/yiliyili/</th>
<th>FREE-V</th>
<th>ALIGN</th>
<th>PARSE</th>
<th>NoCoda</th>
</tr>
</thead>
<tbody>
<tr>
<td>yi.li.yil.&lt;i&gt;</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>yi.li.yi.li</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parse and NoCoda are violated in the first candidate because the final vowel is unparsed and the last consonant is syllabified as a coda. These violations serve to satisfy Free-V, which demands that the word-final vowel be unparsed. Satisfaction of Free-V also forces a violation of the constraint Align, which requires the edge of the word to coincide with a syllable boundary. Prince and Smolensky make two comments about this tableau. The first is on p. 144:

\[(22)\] The relative harmonies of .yi.li.yil.<i> (183 A.i) and .yi.li.yi.li. (183 A.ii) pointedly illustrate the strictness of strict domination. Fully parsed .yi.li.yi.li. is less harmonic than truncated .yi.li.yil.<i> even though it violates only one constraint, while the truncated form violates three of the four lower ranked constraints...

The second is on p. 148:

\[(23)\] **Strictness of strict domination.** In several examples the correct analysis violates many constraints, and its optimality rests crucially on the fact that competitors with a cleaner record overall happen to violate some single dominant constraint. Recall the discussion of /yiliyili/ in 7.3.2: a strong
contender violating just one constraint is bested by an optimal parse violating three of the four less dominant constraints. This effect highlights the content of the central evaluative hypothesis, and sets the theory apart from others in which richer notions of ‘weighting’ and ‘trade-off’ are entertained.

It is in fact not clear how the Lardil example is meant to set ranking apart from weighting. First, we can obviously assign a set of weights to the constraints to pick the correct optimum, as shown in (24). So long as the weight of Free-V is greater than the summed weights of Align, Parse, and NoCoda, [yi.li.yil.<i>] will emerge as optimal.

(24) A weighting capturing the effect of strict domination in Lardil

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>/yiliyili/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yi.li.yi.li.</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-3</td>
</tr>
<tr>
<td>yi.li.yi.li.</td>
<td>-1</td>
<td></td>
<td></td>
<td>-4</td>
</tr>
</tbody>
</table>

**Discussion point:** For any finite set of data, any set of OT optima can be made optimal in HG. Over an infinite set, what kind of pattern can OT express that escapes HG?

Less obviously, any gang effect between the three constraints violated by [yi.li.yil.<i>] would be vacuous. This is due to another kind of specific-to-general relationship that obtains between Free-V and both Align and Parse: Any candidate that satisfies Free-V necessarily violates Align and Parse, but not vice versa (Prince & Smolensky, 1993/2004: 7.2.1). Because Free-V satisfaction entails the violation of these constraints, a gang effect that involves Align and/or Parse with NoCoda in blocking deletion would be vacuous. It would be vacuous because of the failure of these constraints to provide a one-to-one trade-off between NoCoda and Free-V. In the absence of this one-to-one trade-off, there would be no occasion for the lower weight of NoCoda than Free-V to show its effect. To put it differently, the sum of the effects of Align and/or Parse with NoCoda in forcing the violation of Free-V would be the same as the effect of NoCoda alone. As this gang effect is vacuous, it does not produce a divergence in the typological predictions of HG and OT.

**4. Asymmetric trade-offs and HG/OT-LC differences**

The asymmetric trade-off requirement imposes inherent restrictions on the kinds of cumulative interactions that HG can express. Here this is illustrated by a comparison with Smolensky’s (2006b) OT with locally conjoined constraints (henceforth OT-LC). This builds on Legendre et al.’s (2006) comparison of OT-LC and HG, which comes to the opposite conclusion about the relative merits of the two frameworks.
Smolensky first proposed OT-LC in a series of unpublished, but much-cited conference presentations. The literature on OT-LC is now quite large: along with the work cited in Smolensky’s chapter, an especially important further source is Ito & Mester’s (2003) monograph. OT-LC is controversial (see e.g. McCarthy 1999, 2003, Padgett 2002 for critiques; I draw on these here).

Smolensky (2006b: 43) defines local conjunction as follows:

(25) From Pater (2009b)

Local conjunction within a domain $\mathcal{D}$

* A & $\mathcal{D}$ * B is violated if and only if a violation of * A and a (distinct) violation of * B both occur within a single domain of type $\mathcal{D}$.

An example:

(26) Local conjunction analysis of final devoicing

<table>
<thead>
<tr>
<th>/bad/</th>
<th>NoCoda &amp; *VoiceObs</th>
<th>IDENT-Voice</th>
<th>NoCoda</th>
<th>*VoiceObs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [bat]</td>
<td>#!</td>
<td>*</td>
<td>⋆</td>
<td>⋆</td>
</tr>
<tr>
<td>b. bad</td>
<td>#!</td>
<td>*</td>
<td>⋆</td>
<td>⋆</td>
</tr>
<tr>
<td>c. pat</td>
<td>#!</td>
<td>**</td>
<td>⋆</td>
<td>⋆</td>
</tr>
<tr>
<td>d. pad</td>
<td>#!</td>
<td>*</td>
<td>⋆</td>
<td>⋆</td>
</tr>
</tbody>
</table>

Demo: bad-LC.txt

Discussion point: Why can’t HG generate the coda devoicing pattern with just the unconjoined constraints?

Legendre et al. (2006) point to a similar case of HG not being able to generate an attested pattern that OT-LC can get (that a ban on [x] cannot be reduced to independent constraints against velars and fricatives, when the inventory contains both velars and non-velar fricatives), and argue that this counts in OT-LC’s favor.

Whether complex markedness constraints like *CodaVoice and *[x] should be reduced to more basic constraints is a matter of some controversy.

The idea that *CodaVoice is the sum of the effects of a constraint against codas and one against voiced obstruents hearkens back in some ways to theories of prosodic licensing (e.g. Ito 1986, Goldsmith 1990, Lombardi 1991, Steriade 1995), which see contextual markedness as the inability of marked prosodic contexts to license marked segments.

Much work in OT has questioned this approach, analyzing contextual markedness as the effect of rather specific, substantively motivated constraints (e.g. Pater 1999, Steriade 1999). The empirical motivation for these alternatives is that contextual markedness displays asymmetries that are not captured by prosodic licensing: the
set of marked contexts is not the same for every marked segment, and markedness relationships between segments can be reversed across contexts (see Barnes 2006 for an extensive recent critique of prosodic licensing theory). Like prosodic licensing, OT-LC also fails to express these asymmetries.

The critiques of OT-LC have brought out two main issues: it does not require cumulative interactions to be *local*, or co-relevant.

Locality:

(27) **OT-LC analysis of long-distance “coda-devoicing”**

<table>
<thead>
<tr>
<th></th>
<th>/balatak/</th>
<th>NoCODA &amp; ( _Wd _d _d )*VOICEObs</th>
<th>IDENT-VOICE</th>
<th>NoCODA &amp; ( _Wd _d _d )*VOICEObs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.</td>
<td>balatak</td>
<td>!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>ii.</td>
<td>[palatak]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>/balata/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.</td>
<td>[balata]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ii.</td>
<td>palata</td>
<td></td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

Co-relevance:

(28) **Cumulative interaction of coda devoicing and place agreement in OT-LC**

<table>
<thead>
<tr>
<th></th>
<th>/balatak/</th>
<th>NoCODA &amp; ( _Wd _d _d )*VOICEObs</th>
<th>IDENT-VOICE</th>
<th>NoCODA &amp; ( _Wd _d _d )*VOICEObs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.</td>
<td>balatak</td>
<td>!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>ii.</td>
<td>[palatak]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>/balata/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.</td>
<td>[balata]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ii.</td>
<td>palata</td>
<td></td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

**Demo:** badma-LC.txt

Markedness reversal:

(29) **Onset devoicing in OT-LC**

<table>
<thead>
<tr>
<th></th>
<th>/bad/</th>
<th>NoCODA &amp; ( _Seg _Vd _d _d )*VOICEObs</th>
<th>IDENT-VOICE</th>
<th>*VOICEObs</th>
<th>NoCODA &amp; ( _Seg _Vd _d _d )*IDENT-VOICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>[pad]</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>bat</td>
<td>!</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>pad</td>
<td></td>
<td>**!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>pat</td>
<td>!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

In all of these cases, the inability of HG to model the unattested systems generated by OT-LC comes down to the lack of asymmetric trade-offs.
As Legendre et al. (2006) point out, in comparison with HG’s linear model (the harmony function is a linear equation), LC is a superlinear theory of constraint interaction. It does not require asymmetric trade-offs to produce cumulative interactions between constraints.

**Discussion points:** What are some more things that OT-LC can do that HG can or cannot do? Is there anything that HG can produce that OT-LC cannot?

There is a fundamental incoherence in Legendre *et al.*’s case against HG: they argue that it is not powerful enough in comparison with OT-LC, but also that it is too powerful in comparison with plain OT (using a stress example that we’ll talk about next time). In fact, the relatively restrictive theory of cumulative interaction provided by HG opens a number of interesting possibilities for analysis of individual languages and typological study, which we’ll also talk about next time.

**References**


