Cumulative constraint interaction and the equalizer of HG and OT*

Anna Mai and Eric Baković
University of California, San Diego

1 Introduction

In this paper we demonstrate that, in general, Optimality Theory (OT) grammars containing particular, identifiable members of a restricted family of conjoined constraints (Smolensky, 2006) make the same typological predictions as corresponding Harmonic Grammar (HG) grammars. Building on an example case, we propose a general method for identifying the members of this restricted family of conjoined constraints in the *equalizer* of HG and OT, and provide a proof of its intended function. This demonstration adds more structure to claims about the (non)equivalence of HG and OT with local conjunction (Legendre et al., 2006; Pater, 2016) and provides a tool for understanding how different sets of constraints lead to the same typological predictions in HG and OT (Pater, 2016; Jesney, 2016).

The paper is organized as follows. In §2 we provide some background on how OT and HG make different typological predictions in general. We then discuss a relatively simple type of example illustrating how and why those predictions differ in §3, as well as how a conjoined constraint can be identified and added to the existing constraint set to equalize those predictions. A precise algorithm for identifying sets of equalizing conjoined constraints in the general case is provided and discussed in §4, along with a formal proof that the resulting constraint sets are in the equalizer of OT and HG. We conclude in §5 with a summary and prospect.

2 Background

Both OT and HG involve comparisons between different outputs for a given input — different input-output mappings, or *candidates* — and violable constraints. The key distinction between the two frameworks concerns how constraints relate to each other, and therefore the conditions under which constraints are violated by grammatical surface forms (= the *optimal candidates*).[1] In OT, constraints are *ranked* in a strict dominance hierarchy such that the optimal candidate in every candidate set is the candidate that, in every pairwise comparison with another candidate, better satisfies the highest-ranked constraint on which the two candidates differ. In HG, constraints are *weighted* rather than ranked as they are in OT. Violations of a given constraint by a candidate, represented as a negative integer, are multiplied by that constraint's specified (and positive) weight. The sum of the weighted constraint violations incurred by a given candidate is that candidate's *Harmony*, and the optimal candidate is the output candidate with the highest Harmony value.

In this context, a *typology* is a set of grammars (= sets of input-output mappings) predicted of a representative set of candidate sets by a given hypothesis about the content of the constraint set. As Pater (2016: 21) notes, citing Prince & Smolensky (2004: Ch. 10), the difference between constraint ranking and constraint weighting means that the typology predicted in OT will in general be a subset of the typology predicted in HG, given the same sets of constraints and candidate sets. Our goal in this paper is to define a procedure by which, given an arbitrary set of constraints CON, a set of candidate sets *csets*, and the set of optimal candidates in *csets* predicted by CON in OT vs. those predicted in HG, CON is augmented to CON+ such that the predicted OT and HG typologies are the same. The focal point for our demonstration is a relatively simple type of example that we call here the *contrast/neutralization typology*, described in §3.

* Thanks to the members of UC San Diego's PhonCo, the AMP 2019 poster session audience, and Brad Theilman for questions, comments, and suggestions on this work. Work was supported by NIMH training fellowship T32MH020002 and William Orr Dingwall Dissertation Fellowship to A.M. Remaining errors are ours.

[1] The remainder of this section borrows liberally from Baković (2017: 652–653).

3 The contrast/neutralization typology

The types of constraints necessary to describe different patterns of contrast and neutralization in OT are familiar from McCarthy & Prince (1995, 1999): (a) a *specific markedness* constraint, violated by one set of members of an opposition in a specific context; (b) a *general markedness* constraint, violated by the complementary set of members of the opposition more generally; and (c) a *symmetrical faithfulness* constraint, violated by changes from either side of the opposition to the other. The idealized instantiations of these constraint types that we focus on in this paper are given in (1), following Carroll (2012).

(1) Contrast/neutralization constraints (McCarthy & Prince, 1995, 1999; Carroll, 2012)
  a. *Specific markedness*: **si* violated by [si] sequences
  b. *General markedness*: **ʃ* violated by [ʃ]
  c. *Symmetrical faithfulness*: **s*↔*ʃ* violated by /s/ ↦ [ʃ] and /ʃ/ ↦ [s]

Consider a set of four inputs, {/si/, /ʃi/, /sa/, /ʃa/}, representing potential contrasts between *s* and *ʃ* in both the 'palatalizing' context (= __ *i*) and elsewhere (= __ *a*). For each input, there are two relevant candidate outputs: one with [s] and another with [ʃ], in the same context as in the input. These four candidate sets (= *csets*) are spelled out in (2), with the unfaithful segment mappings in each case indicated ***thusly***.

(2) Contrast/neutralization csets
  a. Input: /si/ Candidate outputs: {[si], [***ʃ***i]}
  b. Input: /ʃi/ Candidate outputs: {[***s***i], [ʃi]}
  c. Input: /sa/ Candidate outputs: {[sa], [***ʃ***a]}
  d. Input: /ʃa/ Candidate outputs: {[***s***i], [ʃa]}

With OT's constraint ranking, the factorial typology of the constraints in (1) over the csets in (2) results in the four familiar patterns described in (3a–d). With HG's constraint weighting, the same four patterns are predicted, plus the additional one in (3e) — a 'reverse neutralization' (**RN**) pattern attested in Gujarati.[2]

(3) Contrast/neutralization typology
  a. **F***ull* **C***ontrast*: the $s \sim ʃ$ contrast is preserved in all contexts.
  b. **C***ontextual* **N***eutralization*: the $s \sim ʃ$ contrast is neutralized to *ʃ* before *i* and preserved elsewhere.
  c. **C***omplementary* **D***istribution*: the $s \sim ʃ$ contrast is neutralized to *ʃ* before *i* and to *s* elsewhere.
  d. **A***bsolute* **N***eutralization*: the $s \sim ʃ$ contrast is neutralized to *s* in all contexts.
  e. **R***everse* **N***eutralization*: the $s \sim ʃ$ contrast is preserved before *i* and neutralized to *s* elsewhere.

The additional **RN** pattern is possible in HG due to *ganging cumulativity*, a class of cumulative constraint interactions in which violations of two or more lower-weighted constraints overcome the violations of a single higher-weighted constraint (Legendre et al., 2006; Jäger & Rosenbach, 2006; Pater, 2016; Shih, 2017). Here, the weight of the specific markedness constraint **si* is greater than those of the general markedness constraint **ʃ* and the symmetrical faithfulness constraint **s*↔*ʃ* individually — ⟦ $w(*si) > w(*ʃ) > w(*s{\leftrightarrow}ʃ)$ ⟧ — but the sum of the weights of **s*↔*ʃ* and **ʃ* is greater than that of **si* — ⟦ $w(*s{\leftrightarrow}ʃ)+w(*ʃ) > w(*si)$ ⟧.[3]

Ganging cumulativity results relatively simply in HG from weight additivity, as just shown. In contrast, ganging cumulativity effects cannot be modeled in OT without the inclusion of *conjoined constraints* (Smolensky, 2006), constraints that are systematically violated by a candidate whenever a specified two (or more) other constraints are violated by that candidate. The strictness of strict domination property of ranking otherwise prevents OT from modeling ganging cumulativity. The entire set of typological predictions of HG in (3) can thus be modeled in OT by adding to the constraint set in (1) a conjoined constraint with the necessary ganging cumulativity effect: **ʃ*&**s*↔*ʃ*, violated only when *both* the general markedness constraint **ʃ* *and* the symmetrical faithfulness constraint **s*↔*ʃ* are violated in the same candidate structure. Specifically, the addition of this conjoined constraint to the constraint set in (1) allows OT to generate the **RN** pattern in (3e) under the following total ranking of the constraints: ⟦ $*ʃ\&*s{\leftrightarrow}ʃ \gg *si \gg *ʃ \gg *s{\leftrightarrow}ʃ$ ⟧.

[2] See Carroll (2012) for references, data, and much relevant discussion.
[3] See the Appendix for more details about the typology in (3) and the ranking/weighting conditions responsible for them.

(4) **RN** pattern with constraint conjunction in OT

/si/	*ʃ&*s↔ʃ	*si	*ʃ	*s↔ʃ
☞ [si]		*		
~ [**ʃ**i]	W	L	W	W

/ʃi/	*ʃ&*s↔ʃ	*si	*ʃ	*s↔ʃ
☞ [ʃi]			*	
~ [**s**i]		W	L	W

/sa/	*ʃ&*s↔ʃ	*si	*ʃ	*s↔ʃ
☞ [sa]				
~ [**s**a]	W		W	W

/ʃa/	*ʃ&*s↔ʃ	*si	*ʃ	*s↔ʃ
☞ [**s**a]				*
~ [ʃa]			W	L

Indeed, the predicted typologies of *both* OT and HG, with the conjoined constraint *ʃ&*s↔ʃ added to the constraint set in (1), are exactly the five patterns in (3). Intuitively, this is because the conjoined constraint brings nothing new to the typological table in HG when added to the set of constraints in (1): its sole intended effect is already instantiated in the HG typology due to ganging cumulativity of the general markedness constraint *ʃ and the symmetrical faithfulness constraint *s↔ʃ via weight additivity. This intuition, as well as the other claims in the preceding paragraphs, are formally justified in the next section.

4 The equalizer of HG and OT

Here we identify the structure of constraint sets in the *equalizer* of HG and OT. In category theory, a branch of mathematics concerned with formalizing mathematical structures and systems of structures, an *equalizer* is a set of arguments for which two or more functions have equal values (Riehl, 2017). If we consider HG and OT as functions that take sets of constraints and sets of candidate sets as arguments and map them to language typologies, then the equalizer of HG and OT would be sets of constraints and sets of candidate sets for which HG and OT make the same typological predictions.

Recall that given an arbitrary set CON of (non-conjoined) constraints and a relevant set of candidate sets, the typology generated by HG, HG(CON), is generally a superset of the typology generated by OT, OT(CON), due to the potential for ganging cumulativity in HG (Prince & Smolensky, 2004; Pater, 2016). However, for constraint sets that include a specific set of conjoined constraints, typologies generated in HG and OT will be equivalent. More precisely, these constraint sets will include those conjunctions of constraints that participate in ganging cumulativity in the HG typology.

To demonstrate that this is the case, first we define what we mean by 'conjoined constraint' in §4.1. Then we show how an arbitrary set of (non-conjoined) constraints CON can be augmented with a specific set of conjoined constraints using the procedure summarized in (6) and detailed in §4.2. In §4.3 we define terms necessary for our final demonstration in §4.4 that, whether their predicted typologies differed for CON, the augmented set of constraints CON+ makes equivalent typological predictions in both HG and OT. Thus, the procedure in (6) generates constraint sets in the equalizer of HG and OT.

4.1 *Defining conjoined constraints* Conjoined constraints are defined by the relationship between the violations they assess and the violations assessed by other constraints in the system. In particular, a conjoined constraint $\mathbb{C}_{joint}$ is a constraint that is violated when there is some domain D in which all of $\mathbb{C}_{joint}$'s conjuncts are violated (Padgett, 2002; Łubowicz, 2005; Smolensky, 2006; Legendre et al., 2006; Pater, 2016). The conjoined constraints responsible for points of equivalence between HG and OT, defined in (5), are a specific instance of this type of constraint, where the domain in which all conjuncts are violated is not restricted to any domain smaller than the whole candidate structure.

(5) **Conjoined constraint.** Given a set of candidate sets *csets*, for a subset G of constraints in constraint set CON, $\mathbb{C}_{joint}$ is a *conjoined constraint* for G and *csets* if and only if a violation of $\mathbb{C}_{joint}$ co-occurs with violations of all constraints in G within the same candidate structure.

In other words, if there is no candidate in any tableau for which $\mathbb{C}_{joint}$ is violated without the violation of all constraints in its gang G, and there is no violation of all constraints in G without an accompanying violation of $\mathbb{C}_{joint}$, then $\mathbb{C}_{joint}$ is a conjoined constraint. Note also that violations of conjuncts in the gang G are not restricted to any domain smaller than the full candidate structure. The specification of the full candidate structure as the violation domain ensures that conjoined constraints behave in the same manner as ganging cumulativity in HG, because cumulative constraint interactions in HG are likewise not restricted to violations that are assessed within a proper subdomain of the candidate structure.[4]

Cumulative constraint interaction arises either when the violations of a higher-weighted constraint are overcome by the violations of multiple lower-weighted constraints or by multiple violations of a single lower-weighted constraint. Jäger & Rosenbach (2006) call these two forms of cumulativity *ganging cumulativity* and *counting cumulativity*, respectively. Furthermore, cases of ganging cumulativity themselves can involve counting cumulativity if the number of violations assessed by a particular constraint in the gang plays a crucial role in determining the optimal candidate. Approximating any form of counting cumulativity in OT using conjoined constraints would require a definition of constraint conjunction different from that given in (5). The procedure outlined in (6), and the demonstration in §4.4 that it yields points of predictive equivalence between OT and HG, thus only considers "constraints which are Boolean at the whole-parse level" (Prince & Smolensky, 2004: 98). That is, the following applies only to constraints that assign either a single violation if they are violated anywhere within the candidate structure or assign no violation to the structure at all.

4.2 *Constructing constraint sets in the equalizer of HG and OT* To construct constraint sets for which OT and HG predict the same typology, we begin with an arbitrary constraint set CON. Then, for each pattern in HG(CON) ((6), line 2), we examine each cset in the pattern ((6), line 3). Because this procedure ultimately considers every candidate output structure in every cset ((6), line 6), the candidates assessed in each cset must be restricted to just those that are capable of winning under some ranking of the constraints (the *contenders*; Riggle 2004), to avoid an infinite search. Although the set of candidates given by GEN is in principle infinite, the vast majority of the candidates it contains are bounded by other candidates and thus cannot be optimal under any ranking or weighting of the constraints.[5] Restricting the candidates considered to the finite set of those capable of winning is thus both practical and necessary (Riggle, 2004, 2009).

For each cset, we then compare its optimum against all non-optimal candidates to identify the highest-weighted *distinguishing constraint* ((6), line 4), defined as "a constraint that distinguishes an optimum from another candidate — that is, on which the optimum and some other candidate have different violation scores" (Pater, 2016: 10). If the highest-weighted distinguishing constraint ($\mathbb{H}_d$) does not prefer the optimal candidate, then some form of cumulativity must play a role in determining the optimal candidate. To identify the constraints involved in the cumulative interaction, for each non-optimal candidate ((6), line 6), we compare its violations to those of the optimal candidate and we create a set G of all constraints that both weigh less than $\mathbb{H}_d$ and prefer the optimal candidate ((6), line 7). We then create a constraint $\mathbb{C}_{joint}$ that acts as the conjunction of the constraints in G, using the definition given in (5) ((6), line 8). This conjoined constraint is added to the original set of constraints CON ((6), line 9), and after all losing candidates have been considered in this way, the fully augmented constraint set, which we can now refer to as CON+, is returned ((6), line 10).

(6) A construction for constraint sets in the equalizer of HG and OT

```
1  def equalize(CON):
2    for pattern in HG(CON):
3      for cset in pattern:
4        ℍ_d = the distinguishing constraint with the highest weight
5        if ℍ_d does not prefer the optimal candidate:
6          for loser in cset:
7            G = {ℂ in CON | loser violates ℂ and w(ℍ_d) > w(ℂ)}
8            ℂ_joint = conjoin(G)
9            CON += ℂ_joint
10   return CON
```

[4] This is why we refrain from calling these conjoined constraints *local* conjunctions, as the 'local' delimiter is meant to invoke the (relative) locality of conjunct violations in domains smaller than the full candidate structure. See Legendre et al. (2006) and Pater (2016) for discussion of the non-equivalence of HG and OT with local conjunction.

[5] This is true of both OT and HG for *simple bounding*, where a candidate is bounded by a single other candidate due to the latter having a proper subset of the violations of the former. It is not always true for *collective bounding*, however, where a candidate is bounded by the combined force of two or more other candidates. See Pater (2016: §1.4) for relevant discussion; for more on the simple vs. collective bounding distinction, see Samek-Lodivici & Prince (1999, 2005).

If instead the highest-weighted distinguishing constraint *does* prefer the optimal candidate, the **for**-loop in (6), lines 6–9, is skipped. If this is the case for all csets in all patterns in the HG typology on CON, then the original constraint set is returned unaltered. In these cases, because ganging cumulativity did not contribute to determining any output forms, the HG and OT typologies on the original constraint set are already equivalent, as will be demonstrated in §4.4. But first, some necessary further definitions.

4.3 *Separation and activity* For a constraint set of size n, there exist at most $n!$ languages in the typology predicted by OT, a unique language for each possible total ranking of constraints. However, not every total ranking generates a unique language, because different rankings can (and in fact often do) coincide in the patterns they generate. Some of these coincidences can be due to the particular content of the constraints in a system; for example, because a given constraint simply fails to *separate* a set of candidates.

(7) **OT separation.** A constraint $\mathbb{C}$ *OT-separates* a set of structures if it is satisfied by some members of the set and violated by others. (Prince & Smolensky, 2004: 98)

Other coincidences are motivated by the relationships among constraints. For example, under some rankings, a higher-ranked portion of the constraint hierarchy may filter the candidate set down to a set of structures that a given lower-ranked constraint is now unable to separate.

(8) **OT activity.** Let $\mathbb{C}$ be a constraint in a constraint hierarchy $\mathbb{CH}$, and let i be an input. Then, $\mathbb{C}$ *is OT-active on* i *in* $\mathbb{CH}$ if $\mathbb{C}$ separates the candidates in GEN(i) which are admitted by the portion of $\mathbb{CH}$ that dominates $\mathbb{C}$. (Prince & Smolensky, 2004: 98)

In other words, a constraint $\mathbb{C}$ is active on a candidate set if $\mathbb{C}$ eliminates candidate structures that were not already eliminated by constraints ranked higher than $\mathbb{C}$. When a constraint $\mathbb{C}$ is not active on a given candidate set, the determination of the optimal candidate in that candidate set is not affected by the inclusion of $\mathbb{C}$ in the constraint hierarchy.

Prince & Smolensky's (2004) definition of constraint activity in (8) uses the definition of candidate set separation in (7) to identify constraints that are uniquely capable of filtering candidates due to their ranking within a constraint hierarchy. However, in a weighted constraint system like HG, separation as defined in (7) is incapable of identifying constraints that are active, for an analogous notion of constraint activity in HG. First, the definition of candidate set separation in HG must capture the fact that, in HG, *sets* of constraints can meaningfully separate candidates, as evidenced in cases of ganging cumulativity. A definition of candidate set separation appropriate for HG is given in (9).

(9) **HG separation.** For a subset G of constraints in a weighted constraint set WCON, G *HG-separates* a set S of structures if there exists a subset u of S such that all constraints in G are violated by all members of u and satisfied by all other members of S.

The definition of HG separation in (9) generalizes the definition of OT separation in (7) from individual constraints to sets of constraints. Still, this definition remains insufficient to determine an analogous definition of constraint activity in HG. Strict ranking in OT ensures that if a constraint eliminates candidate structures that were not eliminated by higher-ranked constraints, it is the only constraint to eliminate those structures by virtue of its position within the constraint hierarchy. With constraint weighting, the relative weight of a constraint does not cleanly guarantee the same result, again due to the possibility of ganging cumulativity. Instead, a working definition of constraint activity in HG requires a more precise definition of candidate set separation that we dub *unique separation*, defined in (10).

(10) **Unique separation.** For a subset G of constraints in a weighted constraint set WCON, G *uniquely separates* a set S of candidate structures iff there exists a subset u of S such that G is the unique set of constraints in WCON that is violated by all members of u and satisfied by all other members of S.

Unique separation (10) requires that the subset of constraints that HG-separates a candidate set be the *only* subset of constraints that HG-separates that set of candidate structures. In this way, unique separation facilitates the identification of groups of constraints that actively contribute to ruling out sets of candidate structures and is central to the definition of HG activity given in (11).

(11) **HG activity.** Let G be a subset of constraints in a weighted constraint set WCON, and let i be an input. Then, G *is HG-active on* i *in* WCON if G uniquely separates the candidates in GEN(i).

4.4 *Demonstration of predictive equivalence* In this subsection we show that the conjoined constraint-augmented constraint sets generated by the procedure given in (6) make equivalent typological predictions in both HG and OT frameworks. We call these augmented constraint sets CON+, and we denote the HG typology on CON+ using the notation HG(CON+) and the OT typology on CON+ using the notation OT(CON+). HG(CON+) and OT(CON+) are both sets; therefore, to show that HG(CON+) and OT(CON+) are equal, it suffices to show that OT(CON+) $\subseteq$ HG(CON+) and that HG(CON+) $\subseteq$ OT(CON+).

We know from Prince & Smolensky (2004: Ch. 10) and Pater (2016: 20-21), among others, that for an arbitrary constraint set CON, OT(CON) $\subseteq$ HG(CON). Therefore, this inequality also holds for CON+. To show that HG(CON+) $\subseteq$ OT(CON+) and complete the proof of equality requires the proof of three intermediate statements. First, we show in (12) that for an arbitrary constraint set CON and its corresponding augmentation CON+ by the procedure given in (6), HG(CON+) is equivalent to HG(CON). Second, we show in (13) that OT(CON) is a subset of OT(CON+). Third, we show in (18) that the languages in HG(CON) that are not in OT(CON) are a subset of the languages in OT(CON+). Finally, building off these results, we show in (19) that HG(CON+) $\subseteq$ OT(CON+).

In order for a subset of constraints G to contribute to defining a language in an HG typology, there must exist a set of structures S such that G uniquely separates S. Given this, to demonstrate that HG(CON+) = HG(CON), we show that conjoined constraints are not active in HG. That is, a conjoined constraint in HG never uniquely separates a set of candidate structures.

(12) **Show: HG(CON+) = HG(CON)**
Proof. Let $\mathbb{C}$ be a conjoined constraint. Assume there exists some set of candidate structures S, such that $G = \{\mathbb{C}\}$ uniquely separates S. Then, there exists a subset u of S, such that G is the unique set of constraints that is violated by u and satisfied by all other members of S. That is, if there exists a constraint set G' such that G' is violated by u and satisfied by all other members of S, then $G' = G$. However, by definition (5), $\mathbb{C}$ is a conjoined constraint if and only if its violation co-occurs with the violation of all constraints in some set G'' of constraints. This implies that $G'' = G$, which is a contradiction. □

Next, we show that OT(CON) is a subset of OT(CON+). We accomplish this by showing that for every conjoined constraint $\mathbb{C}_{joint}$ added to CON by the procedure in (6), there exists a ranking of the constraints CON+ such that $\mathbb{C}_{joint}$ is not OT-active in the constraint hierarchy. In particular, we show that conjoined constraints are not active in constraint hierarchies in which they are dominated by any of their conjuncts.

(13) **Show: OT(CON) $\subseteq$ OT(CON+)**
Proof. Let $\mathbb{C}_i$ be an arbitrary conjunct of the conjoined constraint $\mathbb{C}_{joint}$, and let $\mathbb{CH}$ be a constraint hierarchy in which $\mathbb{C}_i \gg \mathbb{C}_{joint}$. Then, $\mathbb{C}_{joint}$ is OT-active if it OT-separates candidate structures that are admitted by the portion of $\mathbb{CH}$ that dominates $\mathbb{C}_{joint}$. That is, there must exist some candidate structure s, such that s violates $\mathbb{C}_{joint}$, and there does not exist a constraint $\mathbb{D}$, such that $\mathbb{D}$ dominates $\mathbb{C}_{joint}$ and s violates $\mathbb{D}$. However, by definition (5), as a conjoined constraint $\mathbb{C}_{joint}$ is violated only if its higher-ranked $\mathbb{C}_i$ conjunct is violated. Therefore, for all constraint hierarchies in which $\mathbb{C}_{joint}$ is dominated by any of its conjuncts, $\mathbb{C}_{joint}$ is not OT-active. □

Every ranking in which $\mathbb{C}_{joint}$ is dominated by at least one of its conjuncts is thus equivalent to that same ranking but with $\mathbb{C}_{joint}$ removed, which means that the effect of every ranking of CON is replicable with a ranking of CON+ simply by ensuring that $\mathbb{C}_{joint}$ is dominated by at least one of its conjuncts.

Next, we show that the languages in HG(CON) that are not in OT(CON) constitute a subset of the languages in OT(CON+). We denote the languages in HG(CON) but not in OT(CON) as HG(CON)\OT(CON), where HG(CON)\OT(CON) is read as 'HG of CON without OT of CON.' As we already know, the languages in HG(CON)\OT(CON) are those with crucial ganging cumulativity. Therefore, to show that the languages in HG(CON)\OT(CON) are a subset of those in OT(CON+), it is sufficient to show that an arbitrary language with crucial ganging cumulativity in HG(CON) is also a member of OT(CON+).

First, we define a language with crucial ganging cumulativity for both HG and OT.

(14) **HG language with crucial ganging cumulativity.** Let WCON be a set of weighted constraints and let i be an input. WCON is said to *produce an HG language with crucial ganging cumulativity* if

there exists a proper subset of constraints $\mathbb{C}_{set} \subset$ WCON and a constraint $\mathbb{C}_{>} \in$ WCON such that all of the following hold:

a. $\mathbb{C}_{>}$ is the highest-weighted distinguishing constraint on GEN(i),
b. $\mathbb{C}_{>}$ prefers a losing candidate $L \in$ GEN(i) to the optimal candidate $W \in$ GEN(i),
c. $\mathbb{C}_{>}$ has higher weight than each constraint in $\mathbb{C}_{set}$,
d. $\mathbb{C}_{set}$ is HG-active on GEN(i), and
e. the summed weight of all constraints in $\mathbb{C}_{set}$ is greater than the weight of $\mathbb{C}_{>}$.

(15) **OT language with crucial ganging cumulativity.** Let $\mathbb{CH}$ be a constraint hierarchy on a set of constraints $\mathbb{C}$ and an input i. $\mathbb{CH}$ is said to *produce an OT language with crucial ganging cumulativity* if there exists a conjoined constraint $\mathbb{C}_{joint}$ such that $\mathbb{C}_{joint}$ is OT-active on GEN(i).

Next, we define a function f between HG and OT languages with crucial ganging cumulativity. For languages with crucial ganging cumulativity, this function takes a set of weighted constraints WCON that produce a language L and maps it to a constraint hierarchy $\mathbb{CH}$ that also produces L.

(16) f : WCON $\rightarrow \mathbb{CH}$
Let CON be a constraint set assessing a set of candidate structures S, and let WCON be a weighting on CON such that WCON produces an HG language with crucial ganging cumulativity in S. Let $\mathbb{C}_{>}$ be defined for WCON as in (14). Then let CON+ be the set of constraints generated from CON by the procedure given in (6), and define a constraint hierarchy on CON+ in the following way:

a. For constraints in CON $\cap$ CON+, let their ranking be determined by the natural ordering on the weights of the constraints in WCON; e.g. $[\![\, w(\mathbb{X}) > w(\mathbb{Y}) \,]\!] \rightarrow [\![\, \mathbb{X} \gg \mathbb{Y} \,]\!]$.
b. For all constraints $\mathbb{C}_{joint}$ in CON+\CON, rank $\mathbb{C}_{joint}$ immediately above $\mathbb{C}_{>}$.
c. Allow conjoined constraints that have the same $\mathbb{C}_{>}$ to be unranked relative to one another.

The resulting constraint hierarchy will produce an OT language with crucial ganging cumulativity. To prove that a constraint hierarchy $\mathbb{CH}$ in the image of f produces an OT language with crucial ganging cumulativity, we show that there exists a conjoined constraint $\mathbb{C}_{joint}$ in $\mathbb{CH}$ and that $\mathbb{C}_{joint}$ is active in $\mathbb{CH}$. We prove these two conditions separately in (17a) and (17b).

(17) Let $\mathbb{CH}$ be a constraint hierarchy in the image of f, and let CON+ be the set of constraints in $\mathbb{CH}$. Let CON be the set of constraints in the pre-image of $\mathbb{CH}$ of f, let WCON be its weighting, and let $\mathbb{C}_{set}$ and $\mathbb{C}_{>}$ be defined for WCON as in (14).

a. **Show: There exists a conjoined constraint $\mathbb{C}_{joint}$ in $\mathbb{CH}$.**
Proof. By (16), WCON produces an HG language with crucial ganging cumulativity. Then, there exists a proper subset of constraints $\mathbb{C}_{set} \subset$ WCON such that the conditions in (14) hold. If these conditions hold, the condition in ((6), line 5) holds, and a conjunction of the constraints in $\mathbb{C}_{set}$ is added to CON+. Therefore, there exists a conjoined constraint $\mathbb{C}_{joint}$ in $\mathbb{CH}$. □

b. **Show: If $\mathbb{C}_{set}$ is HG-active in WCON, then $\mathbb{C}_{joint}$ is OT-active in $\mathbb{CH}$.**
Proof. To the contrary, let $\mathbb{C}_{set}$ be HG-active in WCON, and assume that $\mathbb{C}_{joint}$ is *not* OT-active in $\mathbb{CH}$. Then there exists a set of constraints G in $\mathbb{CH}$, such that $\mathbb{C}_{joint}$ is dominated by all constraints in G and the violations assessed by $\mathbb{C}_{joint}$ are a subset of the violations assessed by the constraints in G. Moreover, by (16a), for all g in the pre-image $f^{-1}(G)$ in WCON, g weighs more than any constraint in $\mathbb{C}_{set}$, and by (5), the violations assessed by the constraints in $\mathbb{C}_{set}$ are a subset of the violations assessed by the constraints in G. However, if the violations assessed by $\mathbb{C}_{set}$ are a subset of the violations assessed by the constraints in G, then $\mathbb{C}_{set}$ is not the unique set of constraints in WCON that is violated by all members of u. Therefore, $\mathbb{C}_{set}$ does not uniquely separate WCON and is not HG-active, leading to a contradiction. □

Now, we show that an arbitrary language with crucial ganging cumulativity in HG(CON) is also a member of OT(CON+).

(18) **Show: HG(CON)\OT(CON) ⊆ OT(CON+)**
Proof. Let CON be a constraint set assessing a set of candidate structures S, and let L be a language in HG(CON)\OT(CON). Then L is an HG language with crucial ganging cumulativity on S, and there exists a weighting WCON on CON such that WCON produces L. Then, by (17), $f(\text{WCON})$ is a constraint hierarchy on CON+ that produces an OT language with crucial ganging cumulativity that is equivalent to L. □

Finally, we show that HG(CON+) is a subset of OT(CON+).

(19) **Show: HG(CON+) ⊆ OT(CON+)**
Proof. We begin with the result demonstrated in (18):

$$\text{HG(CON)}\backslash\text{OT(CON)} \subseteq \text{OT(CON+)}.$$

We then take the union on both sides of this inequality with OT(CON), so that

$$\text{HG(CON)}\backslash\text{OT(CON)} \cup \text{OT(CON)} \subseteq \text{OT(CON+)} \cup \text{OT(CON)}.$$

This implies:

$$\text{HG(CON)} \subseteq \text{OT(CON+)} \cup \text{OT(CON)}.$$

By the result shown in (13), the statement can be further simplified to HG(CON) ⊆ OT(CON+), and by the result shown in (12), HG(CON) = HG(CON+). Therefore, HG(CON+) ⊆ OT(CON+). □

5 Discussion

In the preceding sections we have shown that the generally divergent typological predictions of HG and OT can be made convergent by the targeted addition of conjoined constraints to the constraint set, specifically conjoined constraints the conjuncts of which form a crucial ganging cumulativity set in the HG typology.

Adding conjoined constraints to OT increases the theory's expressivity, allowing it to capture patterns that require cumulative constraint interaction. However, with increased expressivity comes vulnerability to overgeneration. A number of proposals attempt to limit the expressivity of constraint conjunction by placing restrictions on the types of constraints that are permitted to be conjoined with one another (Kirchner, 1996; Fukazawa & Miglio, 1998; Baković, 1999, 2000; Łubowicz, 2005). Other work has shown that the predictions of HG with weight additivity and OT with *local* conjunction diverge (i.e., "Ban Only the Worst of the Worst" (BOWOW) patterns and superadditivity; Smolensky 2006; Padgett 2002; Legendre et al. 2006; Shih 2017). Our results contribute to this literature by showing that the typological consequences of conjoined constraints are structured, and that they bear a principled relationship to weight additivity in HG.

Our results also serve as a useful tool for the probing of different yet related constraint set hypotheses. For example, consider the conjoined constraint *ʃ&*s↔ʃ that results from applying (6) to the HG typology on the constraints in (1). The addition of *ʃ&*s↔ʃ to the constraint set in (1) effectively renders faithfulness *partly asymmetrical*, specifically penalizing mappings from /s/ to [ʃ]. We can thus confidently re-dub this constraint *s→ʃ. The original, symmetrical faithfulness constraint *s↔ʃ is *more stringent* (Prince, 1997) than *s→ʃ, penalizing mappings from /s/ to [ʃ] *and* mappings from /ʃ/ to [s]. But in this case it happens that also adding the other asymmetric faithfulness constraint *ʃ→s to the set {(1), *s→ʃ}, or replacing symmetrical *s↔ʃ with the asymmetrical pair {*s→ʃ, *ʃ→s} altogether, has no additional effect on either the HG typology or the OT typology. Thus the application of (6) to the HG contrast/neutralization typology indirectly reveals that asymmetrical faithfulness constraints might have an advantage over symmetrical faithfulness constraints (if **RN**-type patterns indeed exist, as Carroll (2012) has claimed for Gujarati; recall fn. 3).[6]

In a related vein, Jesney (2016) has shown that where OT requires both positional markedness (PM) and positional faithfulness (PF) constraints to describe both patterns of *disjunctive licensing* (e.g., 'obstruent voicing is licensed only in initial syllables or in onsets') and *conjunctive licensing* (e.g., 'obstruent voicing is licensed only in initial-syllable onsets'), HG can describe *both* patterns with *either* PM constraints *or* PF

[6] Prince (1998) examines the same sets of constraint-types, notes the same difference in typological predictions, and comes to the opposite conclusion about the relative advantages of asymmetrical and symmetrical faithfulness. Thanks to A. Prince for pointing us to this work and discussing it with us.

constraints. In our terms, HG($\textsc{Con}_{PM}$) = HG($\textsc{Con}_{PF}$) = OT($\textsc{Con}_{PM+PF}$), while OT($\textsc{Con}_{PM}$) and OT($\textsc{Con}_{PF}$) are both proper subsets of the other three typologies.[7] Applying (6) to $\textsc{Con}_{PM}$ we obtain the conjoined constraint-augmented $\textsc{Con}+_{PM}$, and applying (6) to $\textsc{Con}_{PF}$ we obtain the conjoined constraint-augmented $\textsc{Con}+_{PF}$. We then know that OT($\textsc{Con}+_{PM}$) = OT($\textsc{Con}+_{PF}$) = OT($\textsc{Con}_{PM+PF}$). From this we can infer that PM constraints behave like the conjoined constraints of $\textsc{Con}+_{PF}$ and PF constraints behave like the conjoined constraints of $\textsc{Con}+_{PM}$. Further investigation of this inference is a topic of our ongoing research.

Appendix

The set of input-output mappings associated with each of the patterns in the typology in (3), and the OT ranking conditions and HG weighting conditions responsible for them given the original set of three constraints in (1), are summarized in Table 1. (Note that there are no OT ranking conditions possible for the **RN** pattern, because the conjoined constraint is not in this constraint set.) When it is in principle possible for the two lower-weighted constraints to participate in a gang against the third, higher-weighted one — as it is for all but the **FC** pattern — the necessary condition on the summed weight of those two lower-weighted constraints is indicated in the second of the two HG weighting conditions shown.

		/si/	/ʃi/	/sa/	/ʃa/	*OT ranking conditions*	*HG weighting conditions*
(3a)	**FC**	[si]	[ʃi]	[sa]	[ʃa]	⟦ *s↔ʃ ≫ **si*, *ʃ ⟧	⟦ w(*s↔ʃ) > w(**si*), w(*ʃ) ⟧
(3b)	**CN**	[**ʃ**i]	[ʃi]	[sa]	[ʃa]	⟦ **si* ≫ *s↔ʃ ≫ *ʃ ⟧	i. ⟦ w(**si*) > w(*s↔ʃ) > w(*ʃ) ⟧ ii. ⟦ w(**si*) > w(*s↔ʃ)+w(*ʃ) ⟧
(3c)	**CD**	[**ʃ**i]	[ʃi]	[sa]	[**s**a]	⟦ **si* ≫ *ʃ ≫ *s↔ʃ ⟧	i. ⟦ w(**si*) > w(*ʃ) > w(*s↔ʃ) ⟧ ii. ⟦ w(**si*) > w(*s↔ʃ)+w(*ʃ) ⟧
(3d)	**AN**	[si]	[**s**i]	[sa]	[**s**a]	⟦ *ʃ ≫ **si*, *s↔ʃ ⟧	i. ⟦ w(*ʃ) > w(**si*), w(*s↔ʃ) ⟧ ii. ⟦ w(*ʃ) > w(*s↔ʃ)+w(**si*) ⟧
(3e)	**RN**	[si]	[ʃi]	[sa]	[**s**a]	✗	i. ⟦ w(**si*) > w(*ʃ) > w(*s↔ʃ) ⟧ ii. ⟦ w(*s↔ʃ)+w(*ʃ) > w(**si*) ⟧

Table 1: Patterns predicted by the constraints in (1) over the csets in (2), and conditions thereupon

The HG weighting conditions shown here for the **RN** pattern differ from those found by the objective function defined in Potts et al. (2010) and used in OT-Help (Staubs et al., 2010), whereby the sum of all the constraint weights is minimized. OT-Help's solution for the **RN** pattern assigns the integer weights shown in (20a). This set of weights is a minimized instance of the set of weighting conditions shown in (20b).

(20) OT-Help's solution for the **RN** pattern

a. Integer constraint weights

w(**si*) = 2 w(*ʃ) = 2 w(*s↔ʃ) = 1

b. Weighting conditions

i. ⟦ w(**si*), w(*ʃ) > w(*s↔ʃ) ⟧

ii. ⟦ w(*ʃ)+w(*s↔ʃ) > w(**si*) ⟧

iii. ⟦ w(**si*)+w(*s↔ʃ) > w(*ʃ) ⟧

The key difference between the set of weighting conditions in (20) and those listed for the **RN** pattern in Table 1 lies in the relative weights of the two markedness constraints, specific **si* and general *ʃ. In Table 1, w(**si*) > w(*ʃ), while in (20), w(**si*) = w(*ʃ). This difference is responsible for the necessity of the third weighting condition in (20b-iii): **si* has to gang up with *s↔ʃ in order to outweigh *ʃ, whereas this instance of ganging cumulativity isn't necessary in Table 1 because **si* independently outweighs *ʃ.

Consider now how each of these two solutions selects the optimal **RN** candidate in each of the four csets. Let's start with the OT-Help solution in (20), the tableaux for which are shown in (21). The (positive) integer

[7] The predictions of HG($\textsc{Con}_{PM}$)/HG($\textsc{Con}_{PF}$) on the one hand and OT($\textsc{Con}_{PM+PF}$) on the other begin to differ slightly when there are three overlappable privileged positions (e.g., *stressed* initial-syllable onsets); see Jesney (2016: §6.2.3).

constraint weights are listed above each constraint, and constraint violations are indicated with negative integers; violation × weight products are added across the columns to arrive at each candidate's Harmony value $\mathcal{H}$, indicated in the final column. (The winner in each case has the $\mathcal{H}$ value closest to zero.)

(21) **RN** mappings in HG, with OT-Help's solution in (20)

/si/	2 **si*	2 *ʃ	1 *s↔ʃ	$\mathcal{H}$
☞ [si]	–1			–2
[**ʃ**i]		–1	–1	–3

/ʃi/	2 **si*	2 *ʃ	1 *s↔ʃ	$\mathcal{H}$
☞ [ʃi]		–1		–2
[***s***i]	–1		–1	–3

/sa/	2 **si*	2 *ʃ	1 *s↔ʃ	$\mathcal{H}$
☞ [sa]				0
[**ʃ**a]		–1	–1	–3

/ʃa/	2 **si*	2 *ʃ	1 *s↔ʃ	$\mathcal{H}$
☞ [***s***a]			–1	–1
[ʃa]		–1		–2

Given this solution, ganging cumulativity is necessary for both the /si/ ↦ [si] and the /ʃi/ ↦ [ʃi] mappings shown in the top two tableaux: /si/ ↦ [si] (in the top-left tableau) requires the weighting condition with ganging cumulativity shown in (20b-ii), while /ʃi/ ↦ [ʃi] (in the top-right tableau) requires the weighting condition with ganging cumulativity shown in (20b-iii).

Compare this with the solution for the **RN** pattern shown in Table 1, accompanied in (22) with a minimized set of weights consistent with the weighting conditions; the tableaux are shown in (23).[8]

(22) Table 1 solution for the **RN** pattern

a. Integer constraint weights

$w(*si) = 4 \quad w(*ʃ) = 3 \quad w(*s{\leftrightarrow}ʃ) = 2$

b. Weighting conditions

i. ⟦ $w(*si) > w(*ʃ) > w(*s{\leftrightarrow}ʃ)$ ⟧

ii. ⟦ $w(*ʃ) + w(*s{\leftrightarrow}ʃ) > w(*si)$ ⟧

(23) **RN** mappings in HG, with the solution in (22)

/si/	4 **si*	3 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [si]	–1			–4
[**ʃ**i]		–1	–1	–5

/ʃi/	4 **si*	3 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [ʃi]		–1		–3
[***s***i]	–1		–1	–6

/sa/	4 **si*	3 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [sa]				0
[**ʃ**a]		–1	–1	–5

/ʃa/	4 **si*	3 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [***s***a]			–1	–2
[ʃa]		–1		–3

Given this solution, ganging cumulativity is only necessary for the /si/ ↦ [si] in the top-left tableau. This mapping requires the weighting condition with ganging cumulativity shown in (22b-ii), similarly to the OT-Help solution. The mapping /ʃi/ ↦ [ʃi] in the top-right tableau, on the other hand, is now selected simply because the weight of **si* is greater than that of *ʃ; there is no need for ganging cumulativity in this case.

The algorithm in (6) works regardless of which HG solution we start from: whether we add to (1) only the *ʃ&*s↔ʃ conjunction, as justified by the solution in (22)–(23), or both *ʃ&*s↔ʃ and **si*&*s↔ʃ, as justified by the solution in (20)–(21), the resultant constraint set yields the same typological predictions in HG as in OT. In fact, the addition of only the **si*&*s↔ʃ conjunction to (1) is also in the equalizer, because

[8] The weights in OT-Help's solution in (20a) sum to 5, while those here in (22a) sum to 9. The sum of the constraint weights of OT-Help's solution is thus maximally minimized in accordance with Potts et al.'s (2010) objective function.

there is a third possible HG solution for the **RN** pattern. This solution reverses the weights of **si* and *ʃ, as shown in (24)–(25).

(24) Third solution for the **RN** pattern

a. Integer constraint weights

w(**si*) = 3 w(*ʃ) = 4 w(*s↔ʃ) = 2

b. Weighting conditions

i. ⟦ w(*ʃ) > w(**si*) > w(*s↔ʃ) ⟧

ii. ⟦ w(**si*)+w(*s↔ʃ) > w(**si*) ⟧

(25) **RN** mappings in HG, with the solution in (24)

/si/	3 **si*	4 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [si]	–1			–3
[**ʃ**i]		–1	–1	–6

/ʃi/	3 **si*	4 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [ʃi]		–1		–4
[***s***i]	–1		–1	–5

/sa/	3 **si*	4 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [sa]				0
[**ʃ**a]		–1	–1	–6

/ʃa/	3 **si*	4 *ʃ	2 *s↔ʃ	$\mathcal{H}$
☞ [***s***a]			–1	–2
[ʃa]		–1		–4

Given this solution, ganging cumulativity is only necessary for the /ʃi/ ↦ [ʃi] in the top-right tableau. This mapping requires the weighting condition with ganging cumulativity shown in (24b-ii), similarly to the OT-Help solution. The mapping /si/ ↦ [si] in the top-left tableau, on the other hand, is now selected simply because the weight of *ʃ is greater than that of **si*; there is no need for ganging cumulativity in this case.

Thus, the following three constraint sets for the contrast/neutralization typology, all of which predict the same typology of five languages in (3), are in the equalizer of OT and HG.

(26) Three constraint sets for the contrast/neutralization typology in the equalizer of OT and HG

constraints in (1)			*conjoined constraints*		
***si**	***ʃ**	***s↔ʃ**	***ʃ&*s↔ʃ**	***si&*s↔ʃ**	
✓	✓	✓	✓	✓	= (20)–(21)
✓	✓	✓	✓		= (22)–(23)
✓	✓	✓		✓	= (24)–(25)

Regardless of the particular weighting solution selected to describe a language in HG, the algorithm given in (6) will return a constraint set in the equalizer of OT and HG. In the main text, we focus on the solution in (22)–(23) for two reasons. First, this is the only solution for which the corresponding OT solution requires a ranking of the constraints in (1) that matches the natural ordering on the weights of the constraints. This parallelism is shown in (27a), where the weighting conditions are abbreviated with the constraints' minimal integer weights. The lack of parallelism in the other two cases is shown in (27b) for (20)–(21) and in (27c) for (24)–(25).[9]

(27) a. ⟦ *ʃ&*s↔ʃ ≫ **si* ≫ *ʃ ≫ *s↔ʃ ⟧   ≃ ⟦ w(**si*) = 4 > w(*ʃ) = 3 > w(*s↔ʃ) = 2 ⟧

b. ⟦ **si*&*s↔ʃ ≫ *ʃ ≫ **si*, *s↔ʃ ⟧   ≄ ⟦ w(**si*) = 2, w(*ʃ) = 2 > w(*s↔ʃ) = 1 ⟧

c. ⟦ *ʃ&*s↔ʃ ≫ **si* ≫ *ʃ ≫ *s↔ʃ ; **si*&*s↔ʃ ⟧
⟦ **si*&*s↔ʃ ≫ *ʃ ≫ **si*, *s↔ʃ ; *ʃ&*s↔ʃ ⟧   ≄ ⟦ w(*ʃ) = 4 > w(**si*) = 3 > w(*s↔ʃ) = 2 ⟧
⟦ *ʃ&*s↔ʃ ≫ **si* ; **si*&*s↔ʃ ≫ *ʃ ≫ *s↔ʃ ⟧

Second, and somewhat more substantively, the conjoined constraint resulting from the solution in (22)–(23) has a natural interpretation as an asymmetrical faithfulness constraint, as discussed in §5.

[9] The third solution has three separate sets of ranking conditions compatible with it, given that each of the conjoined constraints has different ranking conditions depending on the relative rankings of the non-conjoined constraints.

References

Baković, E. (1999). Assimilation to the unmarked. *University of Pennsylvania Working Papers in Linguistics*, vol. 6, 1–16. ROA-340, Rutgers Optimality Archive, http://roa.rutgers.edu.

Baković, E. (2000). *Harmony, Dominance and Control*. Doctoral dissertation, Rutgers University, New Brunswick, NJ. ROA-360, Rutgers Optimality Archive, http://roa.rutgers.edu.

Baković, E. (2017). Review of J. McCarthy and J. Pater (eds.), *Harmonic Grammar and Harmonic Serialism*. *Phonology* 34:3, 651–658.

Carroll, L. (2012). Contrasts and contexts in palatalization. Unpublished ms., UC San Diego.

Fukazawa, H. & V. Miglio (1998). Restricting conjunction to constraint families. *Proceedings of WECOL 1998*, 102–117.

Jäger, G. & A. Rosenbach (2006). The winner takes it all—almost: Cumulativity in grammatical variation. *Linguistics* 44:5, 937–971.

Jesney, K. (2016). Positional constraints in Optimality Theory and Harmonic Grammar. McCarthy, J. J. & J. Pater (eds.), *Harmonic Grammar and Harmonic Serialism*, Equinox, 176–220.

Kirchner, R. (1996). Synchronic chain shifts in optimality theory. *Linguistic Inquiry* 27:2, 341–350.

Legendre, G., A. Sorace & P. Smolensky (2006). The Optimality Theory–Harmonic Grammar Connection. Smolensky, P. & G. Legendre (eds.), *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, MIT Press, vol. II: Linguistic and Philosophical Implications, 339–402.

Łubowicz, A. (2005). Locality of conjunction. ROA-764, Rutgers Optimality Archive, http://roa.rutgers.edu.

McCarthy, J. J. & A. Prince (1995). Faithfulness and reduplicative identity. Beckman, J., L. Walsh Dickey & S. Urbanczyk (eds.), *Papers in Optimality Theory*, GLSA, Amherst, MA, 249–384.

McCarthy, John J. & A. Prince (1999). Faithfulness and identity in prosodic morphology. Kager, R., H. van der Hulst & W. Zonneveld (eds.), *The Prosody-Morphology Interface*, CUP, 218–309.

Padgett, J. (2002). Constraint conjunction versus grounded constraint subhierarchies in optimality theory. ROA-530, Rutgers Optimality Archive, http://roa.rutgers.edu.

Pater, J. (2016). Universal grammar with weighted constraints. McCarthy, J. J. & J. Pater (eds.), *Harmonic Grammar and Harmonic Serialism*, Equinox, 1–46.

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

Prince, A. (1997). Stringency and anti-paninian hierarchies. LSA Linguistic Institute handout, Cornell University. https://ruccs.rutgers.edu/images/personal-alan-prince/gamma/talks/insthdt2.pdf.

Prince, A. (1998). Current Directions in Optimality Theory. Second of *Two Lectures on Optimality Theory*, https://ruccs.rutgers.edu/images/personal-alan-prince/gamma/kobe-all.pdf.

Prince, A. & P. Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Malden, MA.

Riehl, Emily (2017). *Category theory in context*. Courier Dover Publications.

Riggle, Jason (2004). Contenders and learning. *Proceedings of WCCFL 23*, Cascadilla Press, 101–114. ROA-688, Rutgers Optimality Archive, http://roa.rutgers.edu.

Riggle, Jason (2009). Generating contenders. ROA-1044, Rutgers Optimality Archive, http://roa.rutgers.edu.

Samek-Lodivici, Vieri & Alan Prince (1999). Optima. ROA-363, Rutgers Optimality Archive, http://roa.rutgers.edu.

Samek-Lodivici, Vieri & Alan Prince (2005). Fundamental Properties of Harmonic Bounding. ROA-785, Rutgers Optimality Archive, http://roa.rutgers.edu.

Shih, S. S. (2017). Constraint conjunction in weighted probabilistic grammar. *Phonology* 34:2, 243–268.

Smolensky, P. (2006). Optimality in Phonology II: Harmonic Completeness, Local Constraint Conjunction, and Feature Domain Markedness. Smolensky, P. & G. Legendre (eds.), *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, MIT Press, vol. II: Linguistic and Philosophical Implications, 27–160.

Staubs, Robert, Michael Becker, Christopher Potts, Patrick Pratt, John J. McCarthy & Joe Pater (2010). OT-Help 2.0. Software package.