CONSTRAINT SUMMATION AND THE AXIOMATIC DERIVATION OF HG

□ **Introduction.** *Classical* markedness and faithfulness constraints apply to individual candidates. Yet, the literature has also advocated constraints that instead apply to sets of candidates, such as *Distinctiveness constraints* (DCs; Flemming 2008) and *Optimal Paradigm faithfulness constraints* (OPFCs; McCarthy 2005). As a consequence, these approaches "lift" the classical constraints to sets of candidates by summing them across a candidate set. Is this assumption of constraint summation typologically innocuous? We make three contributions: properly formalize this question; show that the answer is positive for an *additive* model of constraint interaction; show that HG can be derived in a principled way from this additivity axiom.

□ **Constraint summation.** DCs embody a preference for more distinct contrasts among surface forms: they penalize pairs of surface forms whose perceptual distance is below a given threshold. For instance, the distinctiveness constraint MinDist in (1) penalizes the pairs of surface forms ([ata], [ada]) and ([ada], [aⁿda]) but not the pair ([ata], [aⁿda]) because segment pairs ([t], [d]) and ([d], ⁿd]) are less distinct perceptually than ([t], [ⁿd]), where prenasalization enhances the voicing contrast with the voiceless stop (Flemming 2004).

(1)

/ata/, /ada/	MinDist	Ident(voice)	Ident(nas)	*ⁿD	*D	*VTV
[ata], [ata]	*	*				**
[ata], [ada]	*				*	*
☞ [ata], [aⁿda]			*	*	*	*
[ada], [ata]	*	**			*	*
[ada], [ada]	*	*			**	
[ada], [aⁿda]	*	*	*	*	**	
[aⁿda], [ata]		**	*	*	*	*
[aⁿda], [ada]	*	*	*	*	**	
[aⁿda], [aⁿda]	*	*	**	**	**	

OPFCs embody a preference for greater similarity among surface forms belonging to the same morphological paradigm: they penalize pairs of paradigm members that differ along some relevant phonological dimension. For instance, the OPFC IdentOP(length) in (2) penalizes paradigms (b) and (c) because the length of the stem-final vowel is not identical in the two surface forms. However, it does not penalize paradigms (a) and (d), where all the vowels standing in correspondence in the two surface forms have the same length (McCarthy 2005).

(2)

/faʕaːl-a/, /faʕaːl-tu/	*VːCCV	IdentOP(length)	IdentIO(length)
(a) [faʕaːl-a], [faʕaːl-tu]	*		
(b) [faʕaːl-a], [faʕal-tu]		*	*
(c) [faʕal-a], [faʕaːl-tu]	*	*	*
☞ (d) [faʕal-a], [faʕal-tu]			**

DCs and OPFCs are formally very different from classical faithfulness and markedness constraints. In fact, classical constraints assign a number of violations to each individual candidate mapping of an underlying form and a corresponding surface realization. DCs and OPFCs instead compare the surface realizations of multiple candidates. This difference has implications for the architecture of grammar. A classical grammar in the constraint-based literature evaluates the candidates of a single UR at a time. A grammar with DCs or OPFCs instead must evaluate sets of candidates corresponding to multiple URs, as illustrated in tableau (1) for the two URs /ata/ and /ada/ and in tableau (2) for the two URs /faʕaːl-a/ and /faʕaːl-tu/.

But what about the classical constraints that are now mixed up with DCs and OPFCs? Flemming and McCarthy make the natural suggestion that classical faithfulness and markedness constraints be "lifted" to sets of candidates by *summing* their constraint violations across all candidates in a set. For instance, in (1), candidate ([ata], [ata]) violates *VTV twice because the two surface forms in this pair each violate it once. In (2), the winner paradigm ([faʕala], [faʕaltu]) violates the input-output faithfulness constraint IdentIO(length) twice because the two surface forms in this pair each violate it once.

□ **Is constraint summation innocuous?** Tableaux (1)/(2) thus have two novelties: they contain non-classical constraints such as DCs and OPFCs; furthermore, the classical constraints are summed over. Do both novelties contribute to the typological predictions of Flemming's and McCarthy's proposals? In other words, if DCs and OPFCs are ranked/weighted at the bottom, do the classical constraints yield the same winners when they are summed over as when they are used classically for a single UR at the time? Or do the classical constraints make different typological predictions when they are summed over as in (1)/(2)?

To formalize this question, we consider two URs (the extension to more than two URs is straightforward). Let A and B be their individual candidate sets, namely the classical tableaux where classical constraints work as usual. Let $<$ be an order over tuples of constraint violations which extends the notion "smaller than" from numbers to tuples. We denote by $\mathrm{opt}_< A$ and $\mathrm{opt}_< B$ the sets of winner candidates in tableaux A and B, namely the sets of those candidates with the "smallest" tuples of violations. We allow $<$ to be a *partial* order, (as needed for HG; see below) whereby $\mathrm{opt}_< A$ and $\mathrm{opt}_< B$ can contain multiple winners.

Let $A \times B$ be the set of pairs (α, β) of a candidate α in A and a candidate β in B. By Flemming's and McCarthy's constraint summation assumption, a candidate pair (α, β) is represented by the sum $\mathbf{a} + \mathbf{b} = (a_1 + b_1, \ldots, a_n + b_n)$ of the tuples of constraint violations $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ of the two candidates α and β. Tableaux (1)/(2) (without MinDist and IdentOP) illustrate $A \times B$. We denote by $\text{opt}_<(A \times B)$ the set of winner pairs in $A \times B$, namely pairs with the smallest summed tuple of violations.

Our first contribution is that **the typological innocuousness of constraint summation relative to a mode of constraint interaction $<$ can be formalized as the identity (3)**: the two URs considered end up with the same winner candidates if we optimize the product candidate set $A \times B$ relative to the summed constraints (left hand side) or if we optimize the two candidate sets A and B separately (right hand side).

$$(3) \quad \underbrace{\text{opt}_< (A \times B)}_{\text{with constraint summation}} = \underbrace{\text{opt}_< A \times \text{opt}_< B}_{\substack{\text{classical approach} \\ \text{without constraint summation}}}$$

- **Typological innocuousness in OT.** The sum $\mathbf{a} + \mathbf{b}$ carries less information than the two individual tuples of constraint violations $\mathbf{a}$ and $\mathbf{b}$: the individual tuples cannot be reconstructed from their sum. One might thus expect (3) to fail because constraint summation wipes away crucial information. This pessimism is dispelled by an independent result due to Prince (2015): he effectively establishes (3) for the special case where $<$ is OT's lexicographic order. Yet, Prince's reasoning relies on ERCs, a piece of notation tailored to OT. His proof is thus involved because constraint summation does not admit a simple counterpart in ERCs. We show that Prince's result admits the following elementary explanation in terms of violation profiles.

  Suppose by contradiction that the candidate pair $(\widehat{\alpha}, \widehat{\beta})$ is OT optimal in $A \times B$ but that, say, the candidate $\widehat{\alpha}$ is not OT optimal in A. Hence, there exists another candidate α in A that beats $\widehat{\alpha}$: the tuple $\mathbf{a} = (a_1, \ldots, a_n)$ of constraint violations of α is smaller than the tuple $\widehat{\mathbf{a}} = (\widehat{a}_1, \ldots, \widehat{a}_n)$ of $\widehat{\alpha}$, namely $\mathbf{a} < \widehat{\mathbf{a}}$. Suppose (without loss of generality) that OT's lexicographic order $<$ is relative to the ranking $C_1 \gg C_2 \gg \ldots \gg C_n$. Thus, $\mathbf{a} < \widehat{\mathbf{a}}$ means (4): the $k-1$ top ranked constraints do not distinguish between the two candidates while the kth constraint decisively assigns less violations to α than to $\widehat{\alpha}$. By adding the corresponding components $\widehat{b}_1, \ldots, \widehat{b}_{k-1}, \widehat{b}_k$ of the tuple $\widehat{\mathbf{b}}$ of constraint violations of candidate $\widehat{\beta}$ to both sides of (4), we obtain (5), which says that $\mathbf{a} + \widehat{\mathbf{b}} < \widehat{\mathbf{a}} + \widehat{\mathbf{b}}$. The candidate pair $(\alpha, \widehat{\beta})$ thus beats the candidate pair $(\widehat{\alpha}, \widehat{\beta})$, contradicting the assumption that the candidate pair $(\widehat{\alpha}, \widehat{\beta})$ is OT optimal in $A \times B$. The proof of the reverse implication is analogous.

$$(4) \quad \begin{array}{c} a_1 = \widehat{a}_1 \\ \vdots \\ a_{k-1} = \widehat{a}_{k-1} \\ a_k < \widehat{a}_k \end{array} \qquad (5) \quad \begin{array}{c} a_1 + \widehat{b}_1 = \widehat{a}_1 + \widehat{b}_1 \\ \vdots \\ a_{k-1} + \widehat{b}_{k-1} = \widehat{a}_{k-1} + \widehat{b}_{k-1} \\ a_k + \widehat{b}_k < \widehat{a}_k + \widehat{b}_k \end{array}$$

- **Typological innocuousness beyond OT.** Does the typological innocuousness of the constraint summation assumption extend beyond OT? In other words, besides OT's lexicographic order, which other ways $<$ of ordering tuples of constraint violations satisfy the identity (3)? The crucial property of OT's lexicographic order used in our analysis above is that (4) entails (5): if we add the same quantity to both sides of the inequality, the inequality is not affected. Thus, let us say that an arbitrary order $<$ over tuples of constraint violations is *additive* (Anderson & Feil 1988) provided, whenever a tuple $\mathbf{a}$ is smaller than a tuple $\mathbf{b}$ and the same tuple $\mathbf{c}$ is added to both, the sum $\mathbf{a}+\mathbf{c}$ is smaller than the sum $\mathbf{b}+\mathbf{c}$ (i.e., $\mathbf{a} < \mathbf{b}$ entails $\mathbf{a}+\mathbf{c} < \mathbf{b}+\mathbf{c}$). Hence, (4)/(5) say that OT's lexicographic order is additive. Our second contribution is that **the identity (3) holds if and only if $<$ is an additive order**. In other words, additive orders provide necessary and sufficient structure for the typological innocuousness of the constraint summation assumption.

- **Implications for HG.** We consider a *utility* or *harmony function* H which assigns to each tuple $\mathbf{a}$ of constraint violations a number $H(\mathbf{a})$. Tuples of constraint violations can then be ordered based on their utility/harmony, with smaller tuples corresponding to a larger harmony: $\mathbf{a} <_H \mathbf{b}$ iff $H(\mathbf{a}) > H(\mathbf{b})$. Based on the discussion above, we are interested in the resulting order $<_H$ being additive. This requires the harmony $H(\mathbf{a}+\mathbf{b})$ of the sum of two tuples $\mathbf{a}, \mathbf{b}$ to coincide with the sum $H(\mathbf{a})+H(\mathbf{b})$ of their harmonies, namely $H(\mathbf{a}+\mathbf{b}) = H(\mathbf{a})+H(\mathbf{b})$. Crucially, an harmony function H satisfies this condition if and only if there exist n weights $w_1, \ldots, w_n$ such that the harmony value $H(\mathbf{a})$ of any tuple $\mathbf{a}$ of constraint violations $a_1, \ldots, a_n$ is equal to the weighted sum of the violations, namely $H(\mathbf{a}) = w_1 a_1 + \ldots + w_n a_n$. Technically, this statement is the Fundamental Theorem of Linear Algebra (Strang 2016), with the twist that H needs not be homogeneous, because constraint violations are integers. In conclusion, our third contribution is that **HG is that utility-based implementation of constraint-based phonology that yields an additive mode of constraint interaction**. In other words, within a utility-based approach, HG is fully justified axiomatically through the requirement that the mode of constraint interaction be additive as motivated above from the phonological perspective of Dispersion Theory and Optimal Paradigm Uniformity.