Sonority-based Stress in Harmonic Grammar: Nontransitive Conflation in Phonological Hierarchies*

Charlie O'Hara
University of Southern California

1 Introduction

Markedness hierarchies note universal markedness implications—given the markedness hierarchy in (1) (with $A \prec B$ encoding the relation A is less harmonic (more marked) than B), no language will find B more marked than A along the dimension represented in the hierarchy.

(1) *Abstract Markedness Hierarchy*
$A \preceq B \preceq C \preceq D$

In the hierarchy in (1) these markedness distinctions are represented with $\preceq$, rather than $\prec$, to represent that these markedness distinctions can be CONFLATED, or ignored on a language specific basis, though they will never be reversed. Conflation patterns are an important empirical domain for evaluating theories of modeling markedness hierarchies in constraint based grammars—de Lacy (2002); de Lacy (2004); de Lacy (2006) shows that fixed rankings of constraints (Prince & Smolensky 1993/2004) predicted a subset of the conflation patterns predicted by stringently related constraints (Prince 1997; Prince 1999) in Optimality Theory (OT; Prince & Smolensky 1993/2004; McCarthy & Prince 1995), with stringently related constraints better matching the actual typology.

Logically, a language could exhibit TRANSITIVE or NONTRANSITIVE conflation. For a conflation pattern to be transitive it is necessary that if A and B are conflated, and B and C are conflated, A and C must be conflated as well. Nontransitive patterns do not maintain this implication.

This paper investigates the conflation patterns predicted using these models in Harmonic Grammar (HG;(Legendre *et al.* 1990; Legendre *et al.* 2006; Pater 2016b)), a version of OT that uses weighted rather than ranked constraints. HG's weighted constraints allow for gang effects, where multiple low weighted violations can outweigh one high weighted violation. Much recent work shows that these gang effects allow patterns to be modeled with a smaller simpler set of constraints (Farris-Trimble 2008; Bane & Riggle 2009; Pater 2012; Potts *et al.* 2010; Jesney 2014; Jesney 2016). This paper shows that nontransitive conflation is an unavoidable consequence in HG—if any conflation is predicted, nontransitive conflation is as well.

*Much thanks to Karen Jesney, Rachel Walker, Reed Blaylock, and Hayeon Jang for feedback on previous drafts and incarnations of this paper. Further thanks to Josh Falk, Brian Hsu, Alessandro Jaker, Paul Kiparsky, Joe Pater, Caitlin Smith, Jochen Trommer and audiences at CLS52 and USC PhonLunch for feedback and helpful discussion and comments on this and related work.

Conflation in sonority based stress, the empirical domain investigated in most of this paper will be introduced in section 2. Section 3 demonstrates how HG differs from OT and models nontransitive conflation. Section 4 shows that in HG the sets of languages predicted by stringently related sets and fixed weightings of constraints are identical. Section 5 shows that nontransitive conflation is a inherent part of modeling markedness hierarchies in HG, regardless of constraint formulization. Finally section 6 offers an example of nontransitive conflation from Japanese (Allen 2013) and describes hypothetical nontransitive conflation patterns on other markedness hierarchies to look for in future research.

2 Sonority-Driven Stress

Sonority-driven stress is exhibited in a language if the choice of stress position can change based on the relative sonority of the vowels in the word. Nganasan[1], a Uralic language, by default shows stress on the penult (2a)[2]. However, if the antepenult is [e, o] or [a], and the penult is a less sonorous vowel than [e o], stress can optionally appear on the antepenult.

(2) *Stress in Nganasan*

a. *Default stress on penult*
   (a) [abˈaʔa] ‘older sister, aunt’
   (b) [imˈiɟi] ‘grandmother’
   (c) [əmkˈətə] ‘from here’

b. *Optional sonority-based stress shift*
   (d) [ɟˈembiʔʃi] ‘dressing’
   (e) [sˈolətu] ‘glass’
   (f) [ˈaniʔə] ‘large’

Importantly, while sonority driven stress shift occurs from [i u ə] to [e o a], there is conflation within those two categories. In (3), stress does not retract from [ə] to high vowels, and it does not retract to [a] from mid vowels.

(3) *Conflation in Nganasan*

a. *No stress shift from Central to High vowels*
   (a) [cintˈəɟi] ‘stoke’
   (b) [cuhˈənu] ‘during’

b. *No stress shift from Mid to Low vowels*
   (c) [bacˈebsa] ‘breathing’
   (d) [lʷamˈobtuʔ] ‘spill, splash’

Sonority based stress patterns, like Nganasan’s rely on the stressed vowel hierarchy, (4), (de Lacy 2002; de Lacy 2004; de Lacy 2006). This hierarchy reflects that less sonorous vowels are more marked when stressed than other vowels.

[1] All Nganasan data is from (de Lacy 2004); de Lacy received the stress description from (Helimski 1998), and gathered additional data from numerous other sources.

[2] If the final syllable is light, which all will be in my data. Heavy final syllables receive stress.

(4) *Stressed Vowel Hierarchy*
ˈə ⪯ ˈi·ˈu ⪯ ˈe·ˈo ⪯ ˈɑ

This hierarchy can be modeled in OT using fixed rankings of constraints or stringently related constraints. Fixed rankings of constraints set the markedness hierarchy by only considering the set of rankings of constraints so that if x is more marked than y, the constraint that marks x dominates y. Under such a system, each relevant constraint is violated by just one tier on the hierarchy as in (5). This model requires a modification to the principle of FACTORIAL TYPOLOGY—any possible ordering of constraints should be a grammatically possible human language. With this set of constraints the markedness hierarchy requires that the rankings in (5c) are maintained in all languages.

(5) a. *HD$_{ft}$/ə, *HD$_{ft}$/i·u, *HD$_{ft}$/e·o, *HD$_{ft}$/ɑ

b. *HD$_{ft}$/x- Incur a violation mark for each vowel that is the head of a foot (stressed vowel) that is an element of x.

c. *Fixed Ranking of constraints*
*HD$_{ft}$/ə ≫ *HD$_{ft}$/i·u ≫ *HD$_{ft}$/e·o ≫ *HD$_{ft}$/ɑ

d. Violations of constraints in a fixed ranking

	*HD$_{ft}$/ə	*HD$_{ft}$/i·u	*HD$_{ft}$/e·o	*HD$_{ft}$/ɑ
a. ˈɑ				-1
b. ˈe			-1	
c. ˈi		-1		
d. ˈə	-1			

A set of stringently related constraints is able to obtain this hierarchy without requiring conditions on the possible rankings. Instead the hierarchies are captured through the definitions of constraints. Where the partitioning constraints in (5) are defined so there is one constraint that is violated by each tier, and each tier violates just one constraint, the stringently defined constraints in (6) are defined so that if any constraint is violated by some tier x, that constraint is also violated by any tier y that is more marked than x. In this case, if a constraint is violated by any stressed vowel x, it is also violated by any less sonorous stressed vowel.

(6) a. *HD$_{ft}$/≤ə, *HD$_{ft}$/≤i·u, *HD$_{ft}$/≤e·o, *HD$_{ft}$/≤ɑ

b. *HD$_{ft}$/≤ x- Assign a violation mark for each vowel that is the head of a foot and is less or equally sonorous to x.

c. *Violations of a stringent set of constraints*

	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤i·u	*HD$_{ft}$/≤e·o	*HD$_{ft}$/≤ɑ
a. ˈɑ				-1
b. ˈe			-1	-1
c. ˈi		-1	-1	-1
d. ˈə	-1	-1	-1	-1

With just the set of stringently related constraints, there is no constraint that favors a more marked tier vs a less marked one—a result not seen with the constraints used in (5). A more marked tier has a superset of the violations of any less marked tier on the stringently related constraints. No matter the ordering of these constraints, a more marked tier on the hierarchy will never be treated as less marked by any grammar.

de Lacy (2002) shows that Nganasan's stress pattern can be modeled in OT only using stringently related constraints. If the penultimate and antepenult have vowels of equal sonority, the only relevant constraints are those that drive default penultimate stress. Here ALIGN-R will be used to prefer penultimate to antepenultimate stress, as in (7).

(7) ALIGN-R- Assign a violation for each syllable between the right edge of the main stressed foot and the right edge of the word.

If two tiers are conflated, all constraints that are violated by the more marked vowel but not the less marked vowel must be ranked below ALIGN-R. In (8), *HD$_{ft}$/≤ə, the only constraint that prefers [ˈi] to [ˈə], is dominated by ALIGN-RIGHT (which favors penultimate position), thus allowing penultimate stress on [ə]. [ˈə] and [ˈi] are treated equally (marked) by all the other sonority-stress constraints. The same basic pattern causes the conflation between [ˈa] and [ˈe], showing that *HD$_{ft}$/≤e·o must be ranked below ALIGN-R.

(8) *Conflation in Nganasan*
ALIGN-R≫*HD$_{ft}$/≤ə, *HD$_{ft}$/≤e·o

/iəV/	*HD$_{ft}$/≤i·u	*HD$_{ft}$/≤ɑ	AL-R	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤e·o
a. (ˈiə)V	*	*	*W	L	*
☞ b. i(ˈəV)	*	*		*	*
/aeV/	*HD$_{ft}$/≤i·u	*HD$_{ft}$/≤ɑ	AL-R	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤e·o
c. (ˈae)V		*	*W		L
☞ d. a(ˈeV)		*			*

In OT with stringently defined constraints, for two tiers to not be conflated, all that is necessary is that one constraint that treats the tiers differently is ranked above ALIGN-R. For adjacent tiers like [ˈi] and [ˈe], this can only be one constraint, in this case *HD$_{ft}$/≤i·u, as in (9).

(9) *HD$_{ft}$/≤i·u≫ALIGN-R

/eiV/	*HD$_{ft}$/≤ɑ	*HD$_{ft}$/≤i·u	AL-R	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤e·o
☞ a. (ˈei)V	*		*		*
b. e(ˈiV)	*	*W	L		*

In OT, this ranking means that there is no conflation between any other tiers which differ on *HD$_{ft}$/≤i·u.

Looking at how conflation works in OT, it becomes apparent that transitivity of conflation patterns is a necessary function of the system. If [ˈə] is conflated with [ˈi] and [ˈi] is conflated with [ˈe], both *HD$_{ft}$/≤ə and *HD$_{ft}$/≤i·u must be ranked below ALIGN-R. Those two constraints are the only ones that differentiate [ˈə] and [ˈe], so those tiers must be conflated as well.

(10) *Conflation is Transitive in OT*

a.

/eiV/	AL-R	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤i·u
☞ a. e(ˈiV)			*
b. (ˈei)V	*W		L

b.

/iəV/	AL-R	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤i·u
☞ a. i(ˈəV)		*	*
b. (ˈiə)V	*W	L	*

c.

/eəV/	AL-R	*HD$_{ft}$/≤ə	*HD$_{ft}$/≤i·u
☞ a. e(ˈəV)		*	*
b. (ˈeə)V	*W	L	L

The conflation pattern in Nganasan is among several that are predicted by stringently defined constraints but not fixed rankings in OT, all shown in (11)

(11) *Conflation in OT (Adaped from (de Lacy 2004))*

Categories				Fixed Ranking	Stringent	Attested
ə	i·u	e·o	ɑ	✓	✓	✓Kobon (Davies 1981)
ə	i·u	e·o	ɑ	✗	✓	✓Gujarati(de Lacy 2002:ch. 3)
ə	i·u	e·o	ɑ	✓	✓	✓Asheninca (Payne 1990)
ə	i·u	e·o	ɑ	✓	✓	✓Yil (Martens & Tuominen 1977)
ə	i·u	e·o	ɑ	✗	✓	?
ə	i·u	e·o	ɑ	✗	✓	✓Nganasan (de Lacy 2004)
ə	i·u	e·o	ɑ	✗	✓	✓Kara (Schlie & Schlie 1993; de Lacy 1997)
ə	i·u	e·o	ɑ	✓	✓	✓many (no sonority based stress)

Note that while the stringently defined set of constraints predicts all the above conflation patterns, the fixed rankings predict a subset—particularly the subset where any conflation between tiers includes [ˈɑ], the least marked tier.

These eight conflation patterns that de Lacy considers are not the only logically possible patterns. Considering that conflation is a relation between two tiers, there are 4 choose 2 (or 6) potential pairs, which each could be conflated or not (giving us 2^6=64 patterns). However, in order to maintain the markedness hierarchy, conflation patterns must be CONTIGUOUS. In order to be contiguous, given a hierarchy $x \preceq y \preceq z$, if x and z are conflated, y must be conflated with both of them as well. Without this we get markedness reversals because either x would be treated as harmonic as z and be less marked than y in this language, or z is treated like x and is more marked than y, both options contradicting the hierarchy.

There remain 6 more contiguous conflation patterns that de Lacy does not consider, shown in the table (12). These patterns are all nontransitive, and thus they cannot be predicted by stringently defined constraints or fixed rankings in OT.

(12) *(Contiguous) Nontransitive Conflation Patterns*

Conflations		
ɑ, e·o	e·o, i·u	ə
ɑ	e·o, i·u	i·u, ə
ɑ, e·o	e·o, i·u	i·u, ə
ɑ, e·o, i·u	i·u, ə	
ɑ, e·o	e·o, i·u, ə	
ɑ, e·o, i·u	e·o, i·u, ə	

3 Nontransitive Conflation in HG

In HG, transitivity is no longer implied by the interaction of stringent constraints. The language shown in (13) shows conflation between [ˈe] and [ˈi], and [ˈi] and [ˈə], but does not show conflation between [ˈe] and [ˈə], so stress shift still occurs between those vowels. Crucially, unlike (10c) in OT, the two lower weighted constraints can gang up to overcome the higher weighted ALIGN-R constraint, as in (13c). Because the constraints are weighted, the sum of $*\text{HD}_{ft}/{\leq}\text{i·u}$ and $*\text{HD}_{ft}/{\leq}\text{ə}$ can surpass ALIGN-R.

(13) *Conflation can be Nontransitive in HG*

a.

	$w = 3$	$w = 2$	$w = 2$	H
/eiV/	AL-R	$*\text{HD}_{ft}/{\leq}\text{ə}$	$*\text{HD}_{ft}/{\leq}\text{i·u}$	
☞ a. e(ˈiV)			-1	-2
b. (ˈei)V	-1W		L	-3

b.

	$w = 3$	$w = 2$	$w = 2$	H
/iəV/	AL-R	$*\text{HD}_{ft}/{\leq}\text{ə}$	$*\text{HD}_{ft}/{\leq}\text{i·u}$	
☞ a. i(ˈəV)		-1	-1	-4
b. (ˈiə)V	-1W	L	-1	-5

c.

	$w = 3$	$w = 2$	$w = 2$	H
/eəV/	AL-R	$*\text{HD}_{ft}/{\leq}\text{ə}$	$*\text{HD}_{ft}/{\leq}\text{i·u}$	
a. e(ˈəV)	W	-1L	-1L	-4
☞ b. (ˈeə)V	-1			-3

In HG, the same weighting conditions are needed to get conflation between adjacent tiers (ALIGN-R over the one constraint that distinguishes the tiers), but for nonadjacent tiers to be conflated, it is necessary that ALIGN-R outweighs the sum of the weights of all constraints that differentiate the tiers, rather than just each constraint individually; i.e. to get conflation between [ˈe] and [ˈə] it is necessary that the sum of $w(*\text{HD}_{ft}/{\leq}\text{ə})$ and $w(*\text{HD}_{ft}/{\leq}\text{i·u})$ is outweighed by ALIGN-R. The tableau in (14a) demonstrates this.

Importantly, this is the only place where weighted constraints expand the typology of conflation patterns. Non-contiguous conflation patterns are still ruled out, without any theoretical limitations on the weightings of the constraints. As an example, a non-contiguous conflation pattern would conflate [ˈə] with [ˈe], but not conflate [ˈi] with either. In order to get conflation between [ˈə] and [ˈe], the sum of $^*\text{HD}_{ft}/{\leq}$i·u and $^*\text{HD}_{ft}/{\leq}$ə must weigh less than ALIGN-R, which implies that both $^*\text{HD}_{ft}/{\leq}$i·u and $^*\text{HD}_{ft}/{\leq}$ə must individually weigh less than ALIGN-R, forcing contiguity.

(14) *Conflation must be Contiguous in HG*

a.

	$w=4$	$w=1$	$w=1$	H
/eəV/	AL-R	$^*\text{HD}_{ft}/{\leq}$ə	$^*\text{HD}_{ft}/{\leq}$i·u	
☞ a. e(ˈəV)		-1	-1	-2
b. (ˈeə)V	-1W	L	L	-4

b.

	$w=4$	$w=1$	$w=1$	H
/eiV/	AL-R	$^*\text{HD}_{ft}/{\leq}$ə	$^*\text{HD}_{ft}/{\leq}$i·u	
a. e(ˈiV)			-1	-1
☜ b. (ˈei)V	-1W		L	-4
/iəV/	AL-R	$^*\text{HD}_{ft}/{\leq}$ə	$^*\text{HD}_{ft}/{\leq}$i·u	
c. i(ˈəV)		-1	-1	-2
☜ d. (ˈiə)V	-1W	L	-1	-5

The weighting conditions for each conflation relations are outlined below:

(15) *Conflation Relation Weighting Conditions*

Conflated Tiers	Weighting Condition for conflation
ə, i	w(ALIGN-R) $>$ w($^*\text{HD}_{ft}/{\leq}$ə)
i, e	w(ALIGN-R) $>$ w($^*\text{HD}_{ft}/{\leq}$i·u)
e, ɑ	w(ALIGN-R) $>$ w($^*\text{HD}_{ft}/{\leq}$e·o)
ə, e	w(ALIGN-R) $>$ w($^*\text{HD}_{ft}/{\leq}$ə)+w($^*\text{HD}_{ft}/{\leq}$i·u)
i, ɑ	w(ALIGN-R) $>$ w($^*\text{HD}_{ft}/{\leq}$i·u) + w($^*\text{HD}_{ft}/{\leq}$e·o)
ə, ɑ	w(ALIGN-R) $>$ w($^*\text{HD}_{ft}/{\leq}$ə) + w($^*\text{HD}_{ft}/{\leq}$i·u) + w($^*\text{HD}_{ft}/{\leq}$e·o)

Thus, Harmonic Grammar with stringently related constraints expands the typology by exactly the six non-transitive contiguous conflation patterns, above in (12). The table in (16) compares the typologies predicted by OT and HG with both stringently defined sets of constraints and fixed rankings/weightings (the following section will show the typology of fixed weightings in HG).

(16) *Possible Conflations with Stringent Constraints*

Conflations				OT		HG	
				Fixed	String.	String.	Fixed
ɑ	e·o	i·u	ə	✓	✓	✓	✓
ɑ	e·o, i·u	ə		✗	✓	✓	✓
ɑ	e·o	i·u, ə		✓	✓	✓	✓
ɑ	e·o, i·u, ə			✓	✓	✓	✓
ɑ, e·o	i·u, ə			✗	✓	✓	✓
ɑ, e·o	i·u	ə		✗	✓	✓	✓
ɑ, e·o, i·u	ə			✗	✓	✓	✓
ɑ, e·o, i·u, ə				✓	✓	✓	✓
ɑ, e·o	e·o, i·u	ə		✗	✗	✓	✓
ɑ	e·o, i·u	i·u, ə		✗	✗	✓	✓
ɑ, e·o	e·o, i·u	i·u, ə		✗	✗	✓	✓
ɑ, e·o, i·u	i·u, ə			✗	✗	✓	✓
ɑ, e·o	e·o, i·u, ə			✗	✗	✓	✓
ɑ, e·o, i·u	e·o, i·u, ə			✗	✗	✓	✓

4 Stringently defined constraints = Fixed Weightings in HG

In OT with strict dominance, fixed rankings are relatively simple: the only sort of parameter to set is whether one constraint dominates another. In HG, though (as well as stochastic OT: see Smith & Moreton (2012)), not only does the linear order of the constraints matters but also the distance between them. This allows fixed weightings (the HG correspondent to fixed rankings) to predict a larger range of possible patterns than fixed rankings in OT.

In order to see if any two tiers are conflated using fixed weightings in HG, we must consider the relative weighting of ALIGN-R and the *markedness difference* between the vowels—the difference between the penalties assigned by the stressed vowel hierarchy constraints (w(*HD$_{ft}$/V$_{penult}$)$-w$(*HD$_{ft}$/V$_{ante}$)).

(17) *Conflation (w(*ALIGN-R*) > w(*HD$_{ft}$)/V$_{penult}$$-w$(*HD$_{ft}$/V$_{ante}$))*

	$w=4$	$w=2$	$w=1$	H
/iəV/	AL-R	*HD$_{ft}$/ə	*HD$_{ft}$/i·u	
☞ a. i('əV)		-1		-2
b. ('iə)V	-1W	L	-1W	-5

(18) *Distinction (w(*HD$_{ft}$/V$_{penult}$)$-w$(*HD$_{ft}$/V$_{ante}$)>w(ALIGN-R))*

	$w=4$	$w=6$	$w=1$	H
/iəV/	AL-R	*HD$_{ft}$/ə	*HD$_{ft}$/i·u	
☞ a. i('əV)		-1		-6
b. ('iə)V	-1W	L	-1W	-5

The markedness difference between the antepenult and the penult being stressed is also calculated in HG with stringently related constraints, and is simply the sum of the weights of all constraints which one vowel violates and the other does not.

Thus, the weighting conditions in (15) can be translated into weighting conditions using the fixed weightings (19).

(19) *Translating Markedness Difference*

Tiers	Stringency	Fixed Weightings
ə, i	*HD$_{ft}$/≤ə	*HD$_{ft}$/ə−*HD$_{ft}$i·u
i, e	*HD$_{ft}$/≤i·u	*HD$_{ft}$/i·u−*HD$_{ft}$/e·o
e, ɑ	*HD$_{ft}$/≤e·o	*HD$_{ft}$/e·o−*HD$_{ft}$/ɑ
ə, e	*HD$_{ft}$/≤ə +*HD$_{ft}$/≤i·u	*HD$_{ft}$/ə−*HD$_{ft}$/e·o
i, ɑ	*HD$_{ft}$/≤i·u + *HD$_{ft}$/≤e·o	*HD$_{ft}$/i·u−*HD$_{ft}$/ɑ
ə, ɑ	*HD$_{ft}$/≤ə + *HD$_{ft}$/≤i·u + *HD$_{ft}$/≤e·o	*HD$_{ft}$/ə−*HD$_{ft}$/ɑ

The typology of conflation patterns can be shown to be the same with stringent versus fixed weightings by showing that for all weights of stringently related constraints, there exists a corresponding fixed weighting, and vice versa. If we let the weight of *HD$_{ft}$/≤ə be the weight of *HD$_{ft}$/ə−*HD$_{ft}$/i·u, and *HD$_{ft}$/≤i·u be *HD$_{ft}$/i·u−*HD$_{ft}$/e·o, we see that *HD$_{ft}$/≤ə +*HD$_{ft}$/≤i·u would equal the difference of *HD$_{ft}$/ə and *HD$_{ft}$/e·o, as predicted above. The same holds for the other sums of stringently related constraints.

There are two important conditions for this, regardless of the constraint set: all constraints must have positive weights; and, for the constraints in the fixed weighting, the fixed weightings are maintained. These conditions enforce each other. Since *HD$_{ft}$/ə must always outweigh *HD$_{ft}$/i·u, their difference will always be positive. Therefore the three constraints *HD$_{ft}$/≤ə , *HD$_{ft}$/≤i·u , and *HD$_{ft}$/≤e·o all will have positive weights. Going the other way, since those constraints have positive weights, the difference between *HD$_{ft}$/ə and *HD$_{ft}$/i·u (and all other differences) must be positive, forcing the fixed weighting to be maintained.

5 Nontransitive Conflation is Unavoidable in HG

Since both stringently defined constraints, and constraints in fixed weightings predict non-transitive conflation, it is prudent to ask if nontransitive conflation can be avoided under any set of constraints in Harmonic Grammar. The answer to this question is no.

Consider any (at least) three tier markedness hierarchy: by definition, Tier-1 must be more marked than Tier-2, and both more marked than Tier-3. These differences in markedness must be finite, because the weights of all constraints in HG are finite.[3] If two tiers are conflated, there must be some constraint whose weight overcomes the markedness difference between them. Thus, if we assume Tier-1 and Tier-2 are conflated, and Tier-2 and Tier-3 are conflated, there must exist some constraint that outweighs each of those markedness differences.

We know that the markedness difference between Tier-1 and Tier-3 must be equal to the sum of the differences between Tier-1 and Tier-2, and Tier-2 and Tier-3; since these are being considered on the one dimensional variable of harmony-score.

[3]This also requires the common assumption that there are only a finite number of constraints in CON. An infinite number of constraints marking Tier-1 and not Tier-2 could also make the markedness difference infinite.

In order to rule out nontransitive conflation, there must be no relevant constraints that have a weight greater than the markedness differences between Tier-1 and Tier-2, and between Tier-2 and Tier-3, but not the difference between Tier-1 and Tier-3. To accomplish this would require an arbitrary restriction to the possible weighting of the other relevant constraints. Importantly, this is not a restriction on the constraints used to model the markedness hierarchy, but any constraints that could interact with the markedness hierarchy.

Potentially, a typology without nontransitive conflation in HG could be modeled with stringently related constraints, by saying that if a constraint outweighs the *Tier≤1 and the *Tier≤2 constraints, it must outweigh the sum of those constraints. However, this must be extended—for any set of markedness hierarchy constraints a constraint outweighs, it must outweigh their sum. This weighting condition doesn't just restrict the weights of other constraints, but it ends up restricting the weights of the constraints within the markedness hierarchy as well. If Tier-1 and Tier-2 are conflated and Tier-2 and Tier-3 are conflated, and Tier-3 and Tier-4 are not conflated, the difference between Tier-3 and Tier-4 must outweigh the difference between Tier-1 and Tier-3. This creates a cascading effect where most constraints are similarly restricted in regards to the sums of most other constraints.

A simple implementation of this restriction could just say that for any constraint C that outweighs the constraints A, B, D, ..., C's weight is greater than the sum of the weights of all those constraints. This restriction effectively turns HG' into OT again, as any gang effects created by interaction of multiple constraints would be banned. Thus the only types of gang effects still created would be those between multiple violations of the same constraint. The vast majority of the gang effects cited to argue for HG's efficacy are of the type that would be banned in this system.

This suggests that nontransitive conflation is a natural and necessary reflex of modeling conflation in HG. Therefore whether nontransitive conflation languages could exist can be used as a strong argument to evaluate whether HG or OT better match the typology of human language.

6 Beginning the Search for Nontransitive Conflation

6.1 Sonority Based Stress - Japanese Pitch Accent

The simplest place to look for nontransitive conflation is sonority based stress systems, as de Lacy did when introducing the concept of conflation. Unfortunately, sonority based stress is rare and conflation patterns can be rare on top of that (Pater 2016b). Identifying nontransitive conflation in sonority based stress requires correctly noting the stress pattern, correctly finding that sonority is involved in stress placement, and finding data showing that adjacent tiers are conflated, but the non-adjacent tiers are not. Further, recent work suggests that Gujarati (one of de Lacy (2004); de Lacy (2006)'s main case studies) does not actually show sonority based stress (Shih to appear), calling sonority based stress in question overall.

Despite all these difficulties, one potential example of nontransitive conflation appears in hiatus-conditioned accent shift in pitch accent words in Japanese (Allen 2013). In Standard Tokyo Japanese, certain verbs have underlying pitch accent, typically on the penultimate mora; but, if no consonant intervenes, the accent can shift to the antepenultimate mora. Allen (2013) proposes a revision to the stressed vowel hierarchy seen in (4), such that front and back vowels are not universally conflated on this tier, as [ˈo] is less marked than [ˈe] and is more sonorous. In (20a), [ao] and [oe] sequences show no accent shift, but [ae] sequences exhibit a shift to the antepenult (20b). (Following Allen (2013)'s notation, accented vowels are marked as V́.)

(20) a. (a) taósu 'to defeat'
(b) koéru 'to become fat'

b. (c) káeru 'to return (intrans.)'
(d) hirugáesu 'to turn over (trans.)'

The actual details of this become more complicated, as many verbs have variable pitch accent shift (NHK 1985; Haraguchi 1996; Allen 2013). However, as a rule [ao] and [oe] sequences shift less often than [ae] sequences, both in real words and Allen's nonce-word study.

As expected, the nontransitive conflation in the Japanese pattern can be modeled in HG using stringent constraints (or a fixed weighting of constraints). In order to find that [á] and [ó], and [ó] and [é] are conflated, we require that the constraints *V́≤o, and *V́≤e, must be weighted below ALIGN-R (or whatever constraint prefers penultimate accents, as shown in (21)).

(21) *Conflation in Japanese*

	$w=3$	$w=2$	$w=2$	H
/taosu/	AL-R	*V́≤o	*V́≤e	
☞ a. taósu		-1		-2
b. táosu	-1W	L		-3
/koeru/	AL-R	*V́≤o	*V́≤e	
☞ c. koéru			-1	-2
d. kóeru	-1W		L	-3

In HG, the two stringently related constraints can gang up to overcome ALIGN-R when both are violated, as in (22), creating nontransitive conflation.

(22) *Conflation in Japanese is nontransitive*

	$w=3$	$w=2$	$w=2$	H
/kaeru/	AL-R	*V́≤o	*V́≤e	
a. kaéru	L	-1W	-1W	-4
☞ b. káeru	-1			-3

In OT, it is not evidently clear how one would model this—it is impossibleusing stringently related constraints (or fixed rankings or any other typical strategy for markedness hierarchies). If [á] and [é] are both conflated with [ó], they must be conflated with each other as well, (23).

(23) *OT cannot model Japanese Conflation*

/taosu/	AL-R	*V́≤o	*V́≤e
☞ a. taósu		*	
b. táosu	*W	L	
/koeru/	AL-R	*V́≤o	*V́≤e
☞ c. koéru			*
d. kóeru	*W		L
/kaeru/	AL-R	*V́≤o	*V́≤e
e. kaéru	L	*W	*W
☜ f. káeru	*		

6.2 Nontransitive Conflation of Place of Articulation

Another domain where nontransitive conflation may be found is along the place of articulation (PoA) markedness hierarchy (24). This section sketches out what a hypothetical language with nontransitive conflation on the place of articulation hierarchy would look like. Cross-linguistic evidence from neutralization and epenthesis creates a hierarchy between consonantal places of articulation (Lombardi 2001; Lombardi 1998; de Lacy 2006).

(24) *Place of Articulation Hierarchy*
dorsal $\preceq$ labial $\preceq$ coronal $\preceq$ glottal

This hierarchy can show conflation much like the stressed vowel hierarchy (de Lacy 2006). For example, in Standard Malay, coronal and glottal are conflated, because /t/ does not debuccalize to [ʔ] in coda (25b), unlike /k/ (25a).

(25) *Malay Debuccalization (from (Lapoliwa 1981) through (de Lacy 2006))*

a. /gərak/→[gə.raʔ] 'move'
(cf. /gərak-an/→[gə.ra.kan]
b. /ikat/→[i.kat] 'to tie'

[t] and [ʔ] must be conflated in Malay, or else debuccalizing /t/ to [ʔ] would be markedness improving. This can be modeled by having ID(PLACE) weighted above *{DORSAL,LABIAL,CORONAL} (but *DORSAL must outweigh the faithfulness).

Further, in Kashaya, dorsal and coronal are conflated, (de Lacy 2006). Kashaya shows debuccalization of all word final plain stops to [ʔ] (26)[4]. However, stops with laryngeal features, whether aspiration or glottalization, fail to debuccalize, (27).

(26) *Kashaya Debuccalization (Buckley 1994; de Lacy 2006)*

a. /qahmat/→ [qahmaʔ] 'angry'
b. /watac/→ [wataʔ] 'frog'
c. /mihjoq/→ [mihjoʔ] 'woodrat'

[4]de Lacy (2006) does not show any examples of debuccalization featuring plain /k/, but shows debuccalization of /t p c q/.

(27) *Kashaya Blocking of Debuccalization*

a. /kilak^h/→[kilak^h] ‘eagle’

b. /hosí́q^ʔ/→[hosí́q^ʔ] ‘ screech owl’

c. /s^{ʔ}ot^ʔ/→[s^{ʔ}ot^ʔ] ‘lungs’

This data shows that dorsal and coronal must be conflated; if they were not, even though the dorsal in /kilak^h/ is blocked from debuccalizing all the way to the glottal stop, it could still reduce in markedness by neutralizing to [kilat^h]. Thus, in Kashaya, *DORSAL (and *{DORS,LAB}) must be outweighed by ID(PLACE).

Nontransitive conflation in a PoA hierarchy would look like a combination of these systems. If a language showed blocking of debuccalization of dorsals like in Kashaya, but only showed debuccalization of plain /k/ and not /t/ like in Standard Malay, it would show nontransitive conflation—dorsal and coronal are conflated, and coronal and glottal are conflated, but dorsal and glottal are not.

(28) *Hypothetical Nontransitive conflation language*

a. /pak/→[paʔ]

b. /tak^h/→[tak^h]

c. /kat/→[kat]

(29) *Nontransitive Conflation language in HG*

	$w = 10$	$w = 3$	$w = 2$	$w = 2$	H
/pak/	*ʔ$^{h/ʔ}$	ID(PLACE)	*DORS	*DORS,COR	
a. pak		L	-1W	-1W	-4
☞ b. paʔ		-1			-3
/tak^h/	*ʔ$^{h/ʔ}$	ID(PLACE)	*DORS	*DORS,COR	
☞ c. tak^h			-1	-1	-4
d. tat^h		-1W	**L**	-1	-5
e. taʔ^h	-1W	-1W	L	L	-13
/kat/	*ʔ$^{h/ʔ}$	ID(PLACE)	*DORS	*DORS,COR	
☞ f. kat				-1	-2
g. kaʔ		-1W		**L**	-3

Crucially, a gang effect is necessary to get this language, both *DORS and *DORS,COR must be outweighed by ID(PLACE) to prevent /tak^h/→[tat^h] and /kat/→[kaʔ] from winning respectively—but the sum of the constraints must surpass ID(PLACE) in order to prevent /pak/→[pak] from winning. This pattern could not be predicted in OT using these constraints, and can only be predicted in HG.

6.3 Conclusion

The stressed vowel hierarchy and the place of articulation hierarchies are just two of the potential hierarchies where nontransitive conflation may appear. Potentially, any phonological hierarchy can have conflation, and any hierarchy with at least three tiers could show a distinction between transitive and nontransitive patterns. The Japanese data presented above offers just one example of nontransitive conflation.

Why have nontransitive conflation systems evaded it us so far? Given that nontransitive conflation seems to be more the norm than the exception in HG, why have we observed so many transitive conflation patterns? There are several potential explanations. First nontransitive systems may have been incorrectly recorded as transitive ones: without enough data it's intuitive to define the conflated tiers as non-overlapping, and therefore transitive. Second, it's possible that we've either missed all the nontransitive languages, or they are an accidental gap in the typology. Third, some non-grammatical factor could be responsible for the gap— Much recent work has investigated the effect of learning on the typology of human languages, including but not limited to Heinz (2009); Heinz (2010); Staubs (2014a); Staubs (2014b); Stanton (to appear); Staubs *et al.* (2016); Hughto *et al.* (2015); Pater (2016a). While all ways of formulating the constraints in HG to achieve a markedness hierarchy predict nontransitive conflation, and stringency and fixed weightings predict the same grammatical typology, the models with different constraint sets have different results for learning, the details of which will be left for future work. Finally, there is the possibility that we are not missing nontransitive conflation, but that nontransitive conflation does not exist. The data from Japanese seems to suggest otherwise, but if there are only a small number of exceptional nontransitive languages, it is possible that grammatical workarounds could be found in OT.

References

Allen, B. 2013. Hiatus-conditioned accent shift in Japanese: expanding the vowel sonority hierarchy. In *UBC Working Papers in Linguistics*, volume 35.

Bane, M., & J. Riggle. 2009. The typological consequences of weighted constraints. In *CLS 45*.

Buckley, E. 1994. *Theoretic aspects of Kashaya phonology and morphology*. CSLI Publications.

Davies, J. 1981. *Kobon*. Amsterdam: North-Holland.

de Lacy, P. 1997. Prosodic categorisation. Master's thesis, University of Aukland.

de Lacy, P. *The formal expression of markedness*. Amherst, MA: University of Massachusetts Amherst dissertation.

de Lacy, P. 2004. Markedness conflation in optimality theory. *Phonology* 21.1–55.

de Lacy, P. 2006. *Markedness: Reduction and Preservation in Phonology*. Cambridge University Press.

Farris-Trimble, A. 2008. Cumulative faithfulness effects: Opaque or transparent? In *IUWPL6: Phonological Opacity Effects in Optimality Theory*, ed. by A. Farris-Trimble & D. A. Dinnsen, 119–145, Bloomington, IN. IULC Publications.

Haraguchi, S. 1996. Syllable, mora and accent. In *Phonological structure and language processing*, ed. by T. Otake & A. Cutler. Mouton de Gruyter.

Heinz, J. 2009. On the role of locality in learning stress patterns. *Phonology* 26.303–351.

Heinz, J. 2010. Learning long-distance phonotactics. *Linguistic Inquiry* 41.623–661.

Helimski, E. 1998. Nganasan. In *The Uralic Languages*, ed. by D. Abondolo. London & New York: Routledge.

Hughto, C., J. Pater, & R. Staubs. 2015. Grammatical agent-based modeling of typology. Paper presented at the GLOW Workshop on Computation, Learnability and Phonological Theory, slides at http://blogs.umass.edu/pater/files/2011/10/hughto-pater-staubs-glow.pdf.

Jesney, K. 2014. Counterbled-counterfeeding in harmonic grammar. In *Proceeding of the 45th Meeting of the North East Lingusitic Society*, ed. by T. Bui & D. Ozyildiz, Amherst, MA. GLSA.

Jesney, K. 2016. Positional constraints in Optimality Theory and Harmonic Grammar. In *Harmonic Grammar and Harmonic Serialism*, ed. by J. J. McCarthy & J. Pater. Equinox.

Lapoliwa, H. 1981. *A generative approach to the phonology of Bahasa Indonesia*. Number 3 in Materials in Languages of Indonesia. Canberra: Pacific Lingusitics Series D-34.

Legendre, G., Y. Miyata, & P. Smolensky. 1990. Harmonic Grammar - a formal multi-level conectionist theory of linguistic wellformedness: an application. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, ed. by L. Erlbaum, 884–891, Cambridge, MA.

Legendre, G., A. Sorace, & P. Smolensky. 2006. The optimality theory-harmonic grammar connection. In *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, ed. by P. Smolensky & G. Legendre, 339–402. MIT Press.

Lombardi, L. 1998. Coronal epenthesis and unmarkedness. In *University of Maryland Working Papers in Linguistics*, volume 5.

Lombardi, L. 2001. Why place and voice are different: Constraint-specific alternations in optimality theory. In *Segmental phonology in Optimality Theory: Constraints and Representations*, ed. by L. Lombardi. Cambridge University Press.

Martens, M., & S. Tuominen. 1977. A tentative phonemic statement of Yil in West Sepik Province. In *Phonologies of five P.N.G. Languages*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

McCarthy, J. J., & A. Prince. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers* 18.249–384.

NHK. 1985. *Nihongo hatsuon akusento jiten [Japanese pronunciation and accent dictionary]*. Tokyo: Nihon Hoosoo Shuppan Kyookai.

Pater, J. 2012. Serial harmonic grammar and berber syllabification. In *Prosody Matters: Essays in Honor of Elisabeth O. Selkirk*, ed. by T. Borowsky, S. Kawahara, T. Shinya, & M. Sugahara, 43–72. Equinox Press.

Pater, J. 2016a. Learning in typological prediction: Grammatical agent-based modeling. Presented at Berkeley Linguistics Society 42.

Pater, J. 2016b. Universal grammar with weighted constraints. In *Harmonic Grammar and Harmonic Serialism*, ed. by J. J. McCarthy & J. Pater. Equinox.

Payne, J. 1990. Asheninca stress patterns. In *Amazonian linguistics: studies in lowland South American languages*, ed. by D. L. Payne, 185–209. Austin: University of Texas Press.

Potts, C., J. Pater, K. Jesney, R. Bhatt, & M. Becker. 2010. Harmonic grammar with linear programming: from linear systems to linguistic typology. *Phonology* 27.77–117.

Prince, A. 1997. Paninian relations. Colloquium Talk, University of Massachusetts Amherst.

Prince, A. 1999. Paninian relations. Handout, University of Marburg.

Prince, A., & P. Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

Schlie, P., & G. Schlie. 1993. A Kara phonology. In *Phonologies of Austronesian Languages 2*, ed. by J. M. Clifton. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Shih, S.-h. to appear. Sonority-driven stress does not exist. In *Supplemental Proceedings of the 2015 Meeting on Phonology*.

Smith, J. L., & E. Moreton. 2012. Sonority variation in Stochastic Optimality Theory: Implications for markedness hierarchies. In *The Sonority Controversy*, ed. by S. Parker, 167–194. Berlin: De Gruyter Mouton.

Stanton, J. to appear. Learnability shapes typology: the case of the midpoint pathology. *Language* .

Staubs, R. *Computational modeling of learning biases in stress typology*. Amherst: University of Massachusetts Amherst dissertation.

Staubs, R. 2014b. Learning and the position of primary stress. In *WCCFL 31*, ed. by R. E. Santana-LaBarge, 428–437.

Staubs, R., J. Culbertson, C. Hughto, & J. Pater. 2016. Grammar and learning in syntactic and phonological typology. Poster Presented at LSA Annual Meeting 2016.