Harmony in Harmonic Grammar by Reevaluating Faithfulness (& Representations)

Charlie O'Hara
University of Southern California

LSA 2016
Washington DC
January 9, 2015

The Problem

- Majority Rules pathologies are a classic issue in OT unbounded harmony. (Lombardi, 1999; Baković, 2000; Finley, 2008)

/+ - - -/	AGREE(F)	ID(F)
a. + + + +		***W
☞ b. - - - -		*
/+ - + +/	AGREE(F)	ID(F)
☞ c. + + + +		*
d. - - - -		***W

- Since [++++] and [- - - -] tie on high ranked markedness, we have to consider the number of ID(F) violations.
- However, with an additional positional faithfulness constraint (Beckman, 1998) weighted above the general faithfulness, this tie can be broken in OT.

The Problem

- Majority Rules pathologies are a classic issue in OT unbounded harmony. (Lombardi, 1999; Baković, 2000; Finley, 2008)

/+ - - -/	AGREE(F)	ID(F)/σ_1	ID(F)
☞ a. + + + +			***
b. - - - -		*W	*L
/+ - + +/	AGREE(F)	ID(F)/σ_1	ID(F)
☞ c. + + + +			*
d. - - - -		*W	***W

- Since [++++] and [- - - -] tie on high ranked markedness, we have to consider the number of ID(F) violations.
- However, with an additional positional faithfulness constraint (Beckman, 1998) weighted above the general faithfulness, this tie can be broken in OT.

The Problem

- With weighted constraints, as in Harmonic Grammar (HG: Legendre *et al.* 1990, 2006; Pater 2009b; Potts *et al.* 2010) we can never ignore the many violations of a low weighted constraint.

	$w = 5$	$w = 3$	$w = 1$	H
/+ - - -/	AGREE(F)	ID(F)/σ_1	ID(F)	H
☞ a. + + + +			-3	-3
b. - - - -		-1	-1	-4
/+ - + +/	AGREE(F)	ID(F)/σ_1	ID(F)	H
☞ c. + + + +			-1	-1
d. - - - -		-1	-3	-6

- Higher weighted constraints like positional faithfulness just act as additional votes against the majority, but with long enough words, the mob rules.

The Problem

- With weighted constraints, as in Harmonic Grammar (HG: Legendre *et al.* 1990, 2006; Pater 2009b; Potts *et al.* 2010) we can never ignore the many violations of a low weighted constraint.

	$w = 5$	$w = 3$	$w = 1$	H
/+ - - -/	AGREE(F)	ID(F)/σ_1	ID(F)	H
☞ a. + + + +			-3	-3
b. - - - -		-1	-1	-4
/+ - - - - -/	AGREE(F)	ID(F)/σ_1	ID(F)	H
c. + + + + + +			-5	-5
☞ d. - - - - - -		-1	-1	-4

- Higher weighted constraints like positional faithfulness just act as additional votes against the majority, but with long enough words, the mob rules.

1. Harmonic Grammar
   - Benefits of weighted constraints
   - Tradeoffs
   - Unbounded Harmony
2. Modifying constraints and representations
   - Markedness Solutions Fail
   - A Faithfulness Solution
   - Issues
3. Conclusion

Intro

- Harmonic Grammar (HG: Legendre *et al.* 1990, 2006; Pater 2009b; Potts *et al.* 2010) is a modification of Optimality Theory (Prince & Smolensky, 1993/2004; McCarthy & Prince, 1995).
- OT uses constraints with a strict ranking.
- HG uses weighted constraints.

Intro

- Harmonic Grammar (HG: Legendre *et al.* 1990, 2006; Pater 2009b; Potts *et al.* 2010) is a modification of Optimality Theory (Prince & Smolensky, 1993/2004; McCarthy & Prince, 1995).
- OT uses constraints with a strict ranking.
- HG uses weighted constraints.

Benefits of weighted constraints

- Allow for language processes to be modeled using fewer and simpler constraints. (Pater 2009a; Potts *et al.* 2010; Pater 2009b, to appear; Jesney 2011, to appear, a.o.)
- Are easily adaptable to handle gradient phenomena. (MaxEnt (Goldwater & Johnson, 2003; Wilson, 2006; Jäger & Rosenbach, 2006) or Noisy HG(Goldrick & Daland, 2009; Boersma & Pater, to appear 2016))
- Offer advantages in language learning (Jesney & Tessier, 2011; O'Hara, 2015)

Bounded Tradeoffs

- Typically in HG, the number of violations incurred by satisfying one violation are bounded. (Pater, 2009a)
- No matter how many voiced obstruents are in a word, the relative weights of ID(VOICE) and *VOICEOBS cause either all or none of them to devoice.
-

	$w = A$	$w = B$	
badagagadabab	*VOICEOBS	ID(VOICE)	HARMONY
a. badagagadabab	-7		-7(A)
b. patakakatapap		-7	-7(B)
$(ba)_n$	*VOICEOBS	ID(VOICE)	HARMONY
c. $(ba)_n$	-n		-n(A)
d. $(pa)_n$		-n	-n(B)

Unbounded Trade-offs

- However, unbounded tradeoffs are a typical issue for harmonic grammar.
- A potentially unbounded number of violations of one constraint can be traded for a single (or bounded) violation of another constraint.

Unbounded Trade-offs

- Legendre *et al.* (2006) show that ALIGN constraints create these effects.

banpata	ALIGNR	SWP	HARMONY
a. 'ban.pa.ta	-2		-2(w(ALIGNR))
b. ban.pa.'ta		-1	-w(SWP)

- We can now set the weights of the constraints so that for any n, only n violations of ALIGNR are tolerated in a word.
  - A language has weight based stress if the heavy syllable is less than 5 syllables from the right edge, but rightmost elsewhere.
  - Or stress lands on heavy syllables within say 400 syllables, but rightmost if all syllables in that window are light.

Unbounded Trade-offs

- Legendre *et al.* (2006) show that ALIGN constraints create these effects.

banpakata	ALIGNR	SWP	HARMONY
a. 'ban.pa.ka.ta	-3		-3(w(ALIGNR))
b. ban.pa.ka.'ta		-1	-w(SWP)

- We can now set the weights of the constraints so that for any n, only n violations of ALIGNR are tolerated in a word.
  - A language has weight based stress if the heavy syllable is less than 5 syllables from the right edge, but rightmost elsewhere.
  - Or stress lands on heavy syllables within say 400 syllables, but rightmost if all syllables in that window are light.

Unbounded Trade-offs

- Legendre *et al.* (2006) show that ALIGN constraints create these effects.

ban$\sigma_n$ta	ALIGNR	SWP	HARMONY
a. 'ban.σ_n.ta	-n		-n(w(ALIGNR))
b. ban.σ_n.'ta		-1	-w(SWP)

- We can now set the weights of the constraints so that for any n, only n violations of ALIGNR are tolerated in a word.
  - A language has weight based stress if the heavy syllable is less than 5 syllables from the right edge, but rightmost elsewhere.
  - Or stress lands on heavy syllables within say 400 syllables, but rightmost if all syllables in that window are light.

Unbounded Trade-offs

- Legendre *et al.* (2006) show that ALIGN constraints create these effects.

ban$\sigma_n$ta	ALIGNR	SWP	HARMONY
a. 'ban.σ_n.ta	-n		-n(w(ALIGNR))
b. ban.σ_n.'ta		-1	-w(SWP)

- We can now set the weights of the constraints so that for any n, only n violations of ALIGNR are tolerated in a word.
  - A language has weight based stress if the heavy syllable is less than 5 syllables from the right edge, but rightmost elsewhere.
  - Or stress lands on heavy syllables within say 400 syllables, but rightmost if all syllables in that window are light.

Unbounded Vowel Harmony

- Tuvan (Turkic) has backness harmony (Harrison, 2000; Rose & Walker, 2011).

(1)	(a)	is-terimden	'footprint' PL-1-ABL.
	(b)	at-tarɯmdan	'name' PL-1-ABL.
	(c)	esker-be-di-m	'notice' NEG-PST.II-1
	(d)	udu-va-dɯ-m	'sleep' NEG-PST.II-1

- The ±back feature from the first vowel in the world spreads to all other vowels.
- Unlike Dominant-Recessive harmony, vowels serve as triggers for harmony based on position not ±back feature value.
- For this analysis I assume that all underlying vowels are specified for the harmonizing feature.

Harmony problem for Harmonic Grammar

- In OT, PosFaith≫GenFaith, causes spreading from privileged positions (positional triggers) and evades Majority Rules pathologies.
-

/is-tar-ɯm/	Agr(Bck)	Id(Bck)/σ_1	Id(Back)
☞ a. isterim			-2
b. ɯstarɯm		-1	-1

Harmony causes problems for Harmonic Grammar

- Since unbounded harmony should work for words of any length, some weighting of PosFaith and GenFaith should avoid majority rules.
- $w(PosFaith) > w(GenFaith)$
-

	$w = 1.5$	$w = 1$	
/is-tar-ɯm/	Id(Back)/Stem	Id(Back)	H
☞ a. isterim		-2	-2
b. ɯstarɯm	-1	-1	-2.5
/is-tar-ɯm-dan/	Id(Back)/Stem	Id(Back)	H
c. isterimden		-3	-3
☜ d. ɯstarɯmdan	-1	-1	-2.5

Harmony causes problems for Harmonic Grammar

- Since unbounded harmony should work for words of any length, some weighting of PosFaith and GenFaith should avoid majority rules.
- $w(PosFaith) > 2 * w(GenFaith)$

	$w = 2.5$	$w = 1$	
/is-tar-ɯm-dan/	Id(Back)/Stem	Id(Back)	H
☞ a. isterimden		-3	-3
b. ɯstarɯmdan	-1	-1	-3.5
/is-tar-ɯm-dan-ar/	Id(Back)/Stem	Id(Back)	H
c. isterimdener		-4	-4
☜ d. ɯstarɯmdanar	-1	-1	-3.5

Harmony causes problems for Harmonic Grammar

- Since unbounded harmony should work for words of any length, some weighting of PosFaith and GenFaith should avoid majority rules.
- $w(PosFaith) > 3 * w(GenFaith)$
-

	$w = 3.5$	$w = 1$	
/is-tar-ɯm-dan-ar/	ID(BACK)/STEM	ID(BACK)	H
☞ a. isterimdener		-4	-4
b. ɯstarɯmdanar	-1	-1	-4.5
/is-tar-ɯm-dan-ar-tɯs/	ID(BACK)/STEM	ID(BACK)	H
c. isterimdenertis		-5	-5
☜ d. ɯstarɯmdanartɯs	-1	-1	-4.5

Harmony causes problems for Harmonic Grammar

- Since unbounded harmony should work for words of any length, some weighting of PosFaith and GenFaith should avoid majority rules.
- $w(PosFaith) > n * w(GenFaith)$
-

	$w = n + .5$	$w = 1$	
/is-dam-(tɯ)$_n$/	ID(BACK)/STEM	ID(BACK)	H
☞ a. isdem(ti)$_n$		-(n+1)	-n-1
b. ɯsdam(tɯ)$_n$	-1	-1	-n-1.5
/isdam-(tɯ)$_{n+1}$/	ID(BACK)/STEM	ID(BACK)	H
c. isdem(ti)$_{n+1}$		-(n+2)	-(n+2)
☜ d. ɯsdem(tɯ)$_{n+1}$	-1	-1	-(n+1.5)

Changing the set of constraints

- The set of constraints we use in HG should not necessarily be the same ones we use in OT, (Jesney, to appear; Pater, to appear).
- Pater (to appear) shows that the unbounded tradeoff created by ALIGNR can be evaded by restricting the types of markedness constraints we have.
  - ALIGNR constraint is problematic because it is gradient.
  - Using categorical constraints (a la McCarthy (2003)) solves this.

CLAIM

However, markedness-based solutions fail to avoid majority rules harmony patterns.

Can't be fixed through markedness

- To drive harmony a markedness constraint M must exist so that $[++_n]$ does better than any non fully harmonic $[+_i-_j]$.
- If a markedness constraint helps prevent majority rules, it must prefer $[++_n]$ to $[--_n]$.

	$w=5$	$w=3$	$w=1$	H
/+ -$_n$/	M	ID(F)/σ_1	ID(F)	H
☞ a. + +$_n$			-n	-n
b. - -$_n$	-n	-1	-1	-(5n+4)
/- +$_n$/	M	ID(F)/σ_1	ID(F)	H
☜ c. + +$_n$		-1	-1	-4
d. - -$_n$	-n		-n	-6n

- But if markedness prefers the expected winner for /+-$_n$/, it must also prefer the expected loser for /-+$_n$/.

Faithfulness

- IDENT(F) is the source of the unbounded tradeoff problem, I will change that constraint.

CLAIM

In harmony processes, features spread rather than just changing.

- Our representations and constraints should distinguish these.

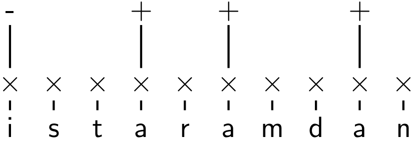
Representational Assumptions

- Features are binary.
- All vowels must be specified for each feature in the input and the output.
- A +F cannot become a -F, it must delete and the -F must be epenthesized.
- For notational simplicity all inputs throughout have no spread features.

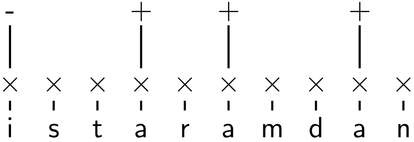
Constraints

- I replace the IDENT(F) family of constraints with the MAX(±F) and DEP(±F) families. (For analyses using these constraints in OT see Lombardi 2001; Walker 1997; Blaho 2008, a.o.)
  - MAX(±F)- Assign a violation mark for any feature ±F in the input with no output correspondent.
  - DEP(±F)- Assign a violation mark for any feature ±F in the output with no input correspondent.
  - MAX(±F)/POS- Assign a violation mark for any feature ±F in the input linked to a segment in POS that has no output correspondent.

(2) *Feature Spreading*

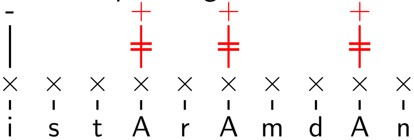


(3) *Feature Changing*

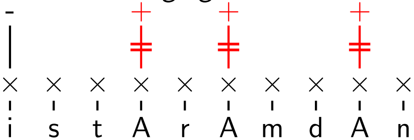


A Faithfulness Solution

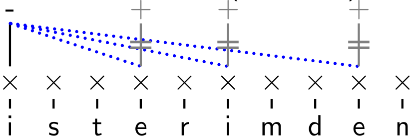
(4) *Feature Spreading*



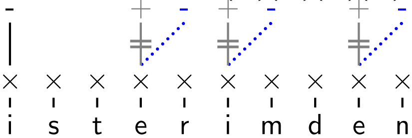
(5) *Feature Changing*



(6) *Feature Spreading* (− _ _ _)



(7) *Feature Changing* (−)(−)(−)(−)



- IDENT F does not care whether features spread or change, but DEP(±F) does.

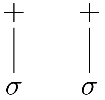
/+−/	ID(BACK)	MAX(±F)	DEP(±F)
a. (+)(+)	-1	-1	-1
b. $(+_{+})$	-1	-1	

(+)(+) $(+_{+})$

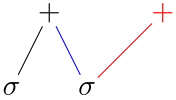
- Further, IDENT F does not differentiate between faithful spreading, and remaining fully faithful, but MAX(±F) does.

/++/	ID(BACK)	MAX(±F)	DEP(±F)
a. (+)(+)			
b. $(+_{+})$		-1	

(+)(+)



$(+_{+})$



- As long as our harmony driving markedness constraint prefers $(+_{+})$ to $(+)(+)$, the majority rules effects will not occur in HG.
  - *Share(F)- Assign a violation mark for any two adjacent vowels that are not linked to the same F feature.

- All fully harmonic candidates to violate MAX(F) for all but one ±F feature.

$+---$	*SHARE(F)
☞ a. $(+_{+++})$	
b. $(_{-}-_{--})$	
c. $(+)(-)(-)(-)$	-3

- Now $[(+_{+++})]$ harmonically bounds $[(_{-}-_{--})]$, preventing majority rules effects.

- All fully harmonic candidates to violate $\textsc{Max}(\mathrm{F})$ for all but one ±F feature.

$+---$	*$\textsc{Share}(\mathrm{F})$	$\textsc{Max}$ F
☞ a. $(+_{+++})$		-3
b. $(_{-}-_{--})$		-3
c. $(+)(-)(-)(-)$	-3	

- Now $[(+_{+++})]$ harmonically bounds $[(_{-}-_{--})]$, preventing majority rules effects.

- All fully harmonic candidates to violate MAX(F) for all but one ±F feature.

$+---$	*SHARE(F)	MAX F	MAX F/σ_1
☞ a. $(+_{+++})$		-3	
b. $(_{-}-_{--})$		-3	-1
c. $(+)(-)(-)(-)$	-3		

- Now $[(+_{+++})]$ harmonically bounds $[(_{-}-_{--})]$, preventing majority rules effects.

- This harmonic bounding prevents majority rules effects at any length.

	$w = 4$	$w = 3$	$w = 1$	
/$+-_n$/	*SHARE(F)	MAX F	MAX F/σ_1	H
☞ a. $(+_{+_n})$		-n		-3n
b. $(-_{-_n})$		-n	-1	-3n-1
c. $(+)(-_{-_{(n-1)}})$	-1	-(n-1)		-3n-1
d. $(+)(-)_n$	-n			-4n

- With these constraints, only candidates a and d can win, depending on the relative weights of MAX F and *A-SPAN(F).

- This harmonic bounding prevents majority rules effects at any length.

	$w = 3$	$w = 4$	$w = 1$	
$/+-_n/$	*SHARE(F)	MAX F	MAX F/σ_1	H
a. $(+_{+_n})$		-n		-4n
b. $(-_{-_n})$		-n	-1	-4n-1
c. $(+)(-_{-_{(n-1)}})$	-1	-(n-1)		-4n+1
☞ d. $(+)(-)_n$	-n			-3n

- With these constraints, only candidates a and d can win, depending on the relative weights of MAX F and *A-SPAN(F).

Issues with this particular solution

- This solution can cause spreading in non-harmony systems to avoid multiple violations of DEP(±F) when adjacent marked features must change.

	$w=3$	$w=1$	$w=5$	
/bytyly/	*y	MAX(±BK)	DEP(±BK)	H
a. (by)(ty)(ly)	-3			-9
b. (bytyly)	-3	-2		-11
c. (bu)(tu)(lu)		-3	-3	-18
☞ d. (butulu)		-3	-1	-8

- This can create an unbounded tradeoff between the segmental markedness constraint and a violation of DEP(F).

$/-_n/$	*-	DEP F
a. $(-)_n$	-n	
☞ b. $(+_n)$		-1

- This suggests we may need to modify how we formulate these markedness constraints as well.

Issues with this particular solution

- This solution can cause spreading in non-harmony systems to avoid multiple violations of DEP(±F) when adjacent marked features must change.

	$w = 3$	$w = 1$	$w = 5$	
/byty/	*y	MAX(±BK)	DEP(±BK)	H
☞ a. (by)(ty)	-2			-6
b. (byty)	-2	-1		-7
c. (bu)(tu)		-2	-2	-12
d. (butu)		-2	-1	-7

- This can create an unbounded tradeoff between the segmental markedness constraint and a violation of DEP(F).

/$-_n$/	*-	DEP F
a. $(-)_n$	-n	
☞ b. $(+_n)$		-1

- This suggests we may need to modify how we formulate these markedness constraints as well.

Results

- Thus, we can see that by modifying our representations and our faithfulness constraints the majority rules effects can be evaded.
- Unbounded tradeoffs are potentially evadable with proper constraints.
    - This can potentially extend to other unbounded tradeoffs, such as those seen with long-distance licensing (Kaplan, 2015), linearity or coda formation (Pater, to appear), etc. (See Bane & Riggle 2009 for more examples of unbounded tradeoffs)

Conclusion

- Unbounded Tradeoff pathologies can be avoided in HG by modifying the set of constraints and representations we are looking at.
- Solutions may cause other problems- but those may be fixable in similar ways.
- If all else fails, the gradualness of serial Harmonic Grammar (Pater, 2012) can also avoid these tradeoffs, but at the expense of adding a new level of complexity to our phonological system.

Thanks

This work developed greatly from discussion with and feedback from Rachel Walker, Karen Jesney, Reed Blaylock, Khalil Iskarous, Louis Goldstein, Caitlin Smith, Jason Riggle, Stephanie Shih, and Jeffrey Heinz. Thanks to participants at USC PhonLunch for feedback on practice versions of this talk and audiences at NELS 46 for feedback on a previous version of this work.

Works Cited I

BAKOVIĆ, ERIC. 2000 (January). *Harmony, dominance and control*. Ph.D. thesis, Rutgers University, New Brunswick, NJ.

BANE, MAX, & RIGGLE, JASON. 2009. The typological consequences of weighted constraints. *In: CLS 45*.

BECKMAN, JILL N. 1998. *Positional Faithfulness*. Ph.D. thesis, University of Massachusetts Amherst, Amherst.

BLAHO, SYLVIA. 2008. *The Syntax of Phonology*. Ph.D. thesis, University of Tromsø.

BOERSMA, PAUL, & PATER, JOE. to appear 2016. Convergence properties of a gradual learning algorithm for Harmonic Grammar. *In:* MCCARTHY, JOHN J., & PATER, JOE (eds), *Harmonic Grammar and Harmonic Serialism*. Equinox.

FINLEY, SARA. 2008 (August). *Formal and Cognitive Restrictions on Vowel Harmony*. Ph.D. thesis, John Hopkins University, Baltimore, Maryland.

GOLDRICK, MATT, & DALAND, ROBERT. 2009. Linking grammatical principles with experimental speech production data: insights from Harmonic Grammar networks. *Phonology*, **26**, 147–185.

GOLDWATER, SHARON, & JOHNSON, MARK. 2003. Learning OT constraint rankings using a Maximum Entropy model. *In: Proceedings of the Workshop on Variation within Optimality Theory*. Stockholm University.

HARRISON, K. DAVID. 2000. *Topics in the Phonology and Morphology of Tuvan*. Ph.D. thesis, Yale University.

JÄGER, GERHARD, & ROSENBACH, ANETTE. 2006. The winner takes it all - almost: cumulativity in grammatical variation. *Linguistics*, **44**, 937–971.

JESNEY, KAREN. 2011. *Cumulative Constraint Interaction In Phonological Acquisition And Typology*. Ph.D. thesis, University of Massachusetts Amherst, Amherst.

JESNEY, KAREN. to appear. Positional Constraints in Optimality Theory and Harmonic Grammar. *In:* MCCARTHY, JOHN J., & PATER, JOE (eds), *Harmonic Grammar and Harmonic Serialism*. Equinox.

JESNEY, KAREN, & TESSIER, ANNE-MICHELLE. 2011. Biases in Harmonic Grammar: The road to restrictive learning. *Natural Language & Linguistic Theory*, **29**.

KAPLAN, AARON. 2015. Long-Distance Licensing in Harmonic Grammar. *In: AMP 2015*.

Works Cited II

Kaun, Abigail. 1995. *The typology of rounding harmony: an optimality theoretic approach*. Ph.D. thesis, UCLA, Los Angeles, CA.

Legendre, Géraldine, Miyata, Yoshiro, & Smolensky, Paul. 1990. Harmonic Grammar - a formal multi-level conectionist theory of linguistic wellformedness: an application. *Pages 884–891 of:* Erlbaum, Lawrence (ed), *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*.

Legendre, Géraldine, Sorace, Antonella, & Smolensky, Paul. 2006. The Optimality Theory-Harmonic Grammar connection. *Pages 339–402 of:* Smolensky, Paul, & Legendre, Géraldine (eds), *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. MIT Press.

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

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

McCarthy, John J. 2003. OT constraints are categorical. *Phonology*, **20**, 75–138.

McCarthy, John J., & Prince, Alan. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers*, **18**, 249–384.

O'Hara, Charlie. 2015. *How abstract is more abstract?: Efficiently searching the set of abstract URs*. M.Phil. thesis, University of Southern California.

Pater, Joe. 2009a. Review of Smolensky and Legendre (2006). The Harmonic Mind. *Phonology*, **26**, 217–226.

Pater, Joe. 2009b. Weighted Constraints in Generative Linguistics. *Cognitive Science*, **33**, 999–1035.

Pater, Joe. 2012. Serial Harmonic Grammar and Berber syllabification. *Pages 43–72 of:* Borowsky, Toni, Kawahara, Shigeto, Shinya, Takahito, & Sugahara, Mariko (eds), *Prosody Matters: Essays in Honor of Elisabeth O. Selkirk*. Equinox Press.

Pater, Joe. to appear. Universal Grammar with Weighted Constraints. *In:* McCarthy, John J., & Pater, Joe (eds), *Harmonic Grammar and Harmonic Serialism*. Equinox.

Works Cited III

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

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

Rose, Sharon, & Walker, Rachel. 2011. Harmony Systems. *In:* van Oostendorp, Marc, Ewen, Colin, Hume, Elizabeth, & Rice, Keren (eds), *Handbook of Phonological Theory*, 2nd ed edn. Wiley-Blackwell.

Steriade, Donca. 1995. Underspecification and markedness. *Pages 114–174 of:* Goldsmith, John A. (ed), *The handbook of phonological theory*, 1 edn. Blackwell.

Walker, Rachel. 1997. Faith and markedness in Esembi feature transfer [ROA-336]. *Pages 103–115 of:* Walker, Rachel, Katayama, Motoko, & Karvonen, Daniel (eds), *Phonology at Santa Cruz*, vol. 5.

Wilson, Colin. 2006. Learning phonology with a substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science*, **30**(5), 945–982.

4 Appendix
- Antilicensing Solution
- Linking Constraints and Underlyingly spread features
- Privative vs. Binary features
- Unspecification of UR

Anti-licensing Solution

- By quantifying violations on the feature, rather than the segment, we can avoid this problem.
- *FroRo (Adapted from Kaun (1995))- Assign a violation mark for each [-back] feature linked to a segment with a [+round] feature.

	$w = 3$	$w = 1$	$w = 1$	
/bytyly/	*FroRo	Max(±bk)	Dep(±bk)	H
a. (by)(ty)(ly)	-3			-9
b. (bytyly)	-1	-2		-5
c. (bu)(tu)(lu)		-3	-3	-6
☞ d. (butulu)		-3	-1	-4

- Now, whether each span backs depends on the relative weighting of *FroRo and Max±bk+Dep(±bk) (potentially including a positional constraint for when privileged constraints are involved.)

- This set of constraints makes overall correct seeming typological predictions, looking just at featural markedness constraints i.e. *+F.
- One issue is no constraint differentiates $(++_m)\text{-}_p$ from $+(\text{-}_m\text{-})\text{-}_{p-1}$

$++_m-_p$	*[+F]	Max F	Dep F	*Share(F)	Max F/σ_1
a. $(++_m)-_p$	-1	-m		-p	
b. $+(-_m-)-_{p-1}$	-1	-m		-p	

Cooccurence Constraints

- How we formulate feature cooccurence constraints is not a trivial question.
- We could have:
  - *FroRo- Assign a violation mark for each [-back] feature linked to a feature that is also linked to a [+round] feature.
  - *RoFro- Assign a violation mark for each [+round] feature linked to a feature that is also linked to a [-back] feature.
- However, feature cooccurence constraints like *FroRo become problematic because you can trade 1 violation of *FroRo for many violations of Max(±round).

	$w = 8$	$w = 2$	$w = 1$	$w = 1$	
/bytyly/	*FroRo	Mx(±rd)	Dep(±rd)	Mx(±bk)	H
a. ([by][ty][ly])	-1			-2	-10
☞ b. [(bitiliti)]		-3	-1	-2	-8

Cooccurence Constraints

- How we formulate feature cooccurence constraints is not a trivial question.
- We could have:
  - *FroRo- Assign a violation mark for each [-back] feature linked to a feature that is also linked to a [+round] feature.
  - *RoFro- Assign a violation mark for each [+round] feature linked to a feature that is also linked to a [-back] feature.
- However, feature cooccurence constraints like *FroRo become problematic because you can trade 1 violation of *FroRo for many violations of Max(±round).

	$w = 8$	$w = 2$	$w = 1$	$w = 1$	
/bytylyty/	*FroRo	Mx(±rd)	Dep(±rd)	Mx(±bk)	H
a. ([by][ty][ly][ty])	-1			-3	-11
☞ b. [(bitiliti)]		-4	-1	-3	-12

Link Constraints

- Typically along with featural Max and Dep, constraints against the linking of features are used as well.
- I did not show my DEPLINK(±F) constraint here, because in all examples it acts the same as MAX(±F).
  - DEPLINK(±F)- Assign a violation mark for any association line between a ±F feature and a segment in the output that has no input correspondent.

$+-$	MAX(±F)	DEPLINK(±F)
a. $(+)(-)$		
b. $(+_+)$	-1	-1
c. $(+)(+)$	-1	-1
d. $(_-)$	-1	-1
$++$	MAX(±F)	DEPLINK(±F)
e. $(+_+)$	-1	-1
f. $(-_)$	-2	-2

Underlying spread features

- These do differ when inputs can have underlying spread features.

$(+)(-_{--})$	MAX(±F)	DEPLINK(±F)
a. $(+)(-_{--})$		
b. $(+_{+++})$	-1	-3
c. $(+)(+)(+)(+)$	-1	-3
d. $({}_{-}-_{--})$	-1	-1
$++$	MAX(±F)	DEPLINK(±F)
e. $(-_{-})$	-2	-2

- This can create unbounded tradeoffs between DEPLINK(±F) and MAX(±F)/σ_1

$(+)(-_{-_n})$	MAX(±F)/σ_1	DEPLINK(±F)
a. $(+_{+_{n+1}})$		-(n+1)
b. $({}_{-}-_{-_n})$	-1	-1

Potential Solution

- By referring to spans rather than links, we can avoid this issue.
- DONTEXTENDSPAN(±F) (DES)- Assign a violation mark for each vowel that is a member of a ±F-span on the output, that was not a member of that span in the input.
- I define *a ±F-span* to be a set of segments that are all linked to the same feature ±F. A member of a span is any segment that is an element of a span.

$(+)(-_{-_n})$	MAX(±F)/σ_1	DES(±F)
a. $(+_{+_{n+1}})$		-1
b. $(_{-}-_{-_n})$	-1	-1

Privative Features

- Convincing arguments have been made that certain features are privative and monovalent. (for one example see Steriade (1995) for nasal)
- I claimed crucially that features like $\pm$back where both feature values can drive harmony are binary.
- These accounts seem to be at odd, but perhaps this is unproblematic.
- Nasal spreading patterns differently than $\pm$back harmony, i.e. we never see [-nasal] spreading.
- Thus, nasal spreading could be driven by a markedness account, that we showed could not work for vowel harmony, since nasal spreading is of the dominant-recessive type.
- Future work will investigate these hypotheses.

Unspecification of URs

- Throughout, I've assumed that all underlying vowels are linked to some ±F feature for vowel place features like back.
- This assumption is not trivial.
- If underlying segments can be unspecified for ±F, but output segments must all be linked to ±F, we further differentiate MAX(±F) from DES(±F).

(+)0	MAX(±F	DES(±F)
a. $(+_+)$		-1
(+)(−)	MAX(±F)	DES(±F)
b. $(+_+)$	-1	-1

- If unprivileged positions are unspecified, there is little typological effect.

Unspecification of URs

- However if the privileged position is unspecified can create strange effects.
- If all unprivileged syllables agree underlyingly, the whole word harmonizes to that feature.

$0(-)_n$	MAX(±F)/σ_1	MAX(±F)	DES(±F)
a. $(+_{+_n})$		-n	-(n+1)
☞ b. $(-_{-_n})$		-(n-1)	-n

- If any disagree, the fully harmonizing candidates tie, so this falls to other constraints (segmental markedness, perhaps otherwise irrelevant positional faithfulness)

$0(-)_n(+)$	MAX(±F)/σ_1	MAX(±F)	DES(±F)
a. $(+_{+_n})$		-n	-(n+1)
☞ b. $(-_{-_n})$		-n	-(n+1)

Underspecification of URs

- This predicts that languages would have positional driven harmony in words with specified initial syllables, but could have dominant recessive harmony or harmony driven from somewhere else when the initial syllable is unspecified.