How powerful is too powerful? Constraint conjunction in weighted constraint grammar and its typological consequences

Aljoša Milenković
Harvard University

Annual Meeting on Phonology 2023
October 21, 2023

Roadmap

Monomoraic feet

- Bosnian/Croatian/Montenegrin/Serbian (BCMS) permits degenerate (=monomoraic) feet

(1)	a.	ˈ(brát)	‘brother.NOM.SG’	(ˈbrá.ta)	‘brother.GEN.SG’
	b.	ˈ(dɛ́d)	‘grandpa.NOM.SG’	(ˈdɛ́.du)	‘grandpa.DAT.SG’
	c.	ˈ(líx)	‘pour.AOR.1SG’	(ˈlí.smɔ)	‘pour.AOR.1PL’

Toneless foot-heading moras

- BCMS tolerates feet headed by a toneless (=non-High-toned) mora

(2)	a.	ˈ(ʋuuk)	‘wolf.NOM.SG’	ˈ(ʋuu).ka	‘wolf.GEN.SG’
	b.	ˈ(graad)	‘city.NOM.SG’	ˈ(graa).da	‘guest.GEN.SG’
	c.	ˈ(rɛɛʧ)	‘word.NOM.SG’	ˈ(rɛɛ).ʧi	‘word.NOM.PL’

No toneless monomoraic feet

- BCMS **categorically prohibits** degenerate feet with a toneless head mora ⇒ GANG EFFECT
- Monosyllabic Lengthening (ML; Zec 1999) in **toneless** (3), but not in High-toned monomoraic forms (1)

(3)	a.	ˈ(lɛɛd)	‘ice.NOM.SG’	(ˈlɛ.da)	‘ice.GEN.SG’
	b.	ˈ(gɔɔst)	‘guest.NOM.SG’	(ˈgɔ.sta)	‘guest.GEN.SG’
	c.	ˈ(stɔɔ)	‘hundred’	ˈ(pɛɛt).stɔ	‘five hundred’

Prosodic minimality in BCMS

(4) Prosodic Minimality in BCMS

a. All bimoraic feet are well-formed.

b. Monomoraic feet must be headed by a High-toned mora.

c. Monomoraic feet with a non-High-toned head mora are **subminimal**.

Conspiracy

- Monosyllabic Lengthening (3) is not the sole strategy employed to eliminate toneless monomoraic feet in BCMS
- Multiple processes conspiring against toneless monomoraic feet across BCMS dialects (see Appendix 1)

Constraints

Markedness:

(5) Head-H (Yip 2001)
Assess a violation for every foot-heading mora not associated with a High tone.

(6) FootBinarity (FtBin) (moraic version) (Prince 1990)
Assess a violation for every monomoraic foot.

Faithfulness:

(7) Dep-μ (McCarthy and Prince 1995)
No mora insertion (= penalizes vowel lengthening)

(8) Dep-H (McCarthy and Prince 1995)
No High tone insertion.

Constraint ranking/strength

- No lengthening in /$\acute{\mu}$/: DEP-μ $\gg$ FTBIN

(9)

brát	DEP-μ	FTBIN
a. ☞ ˈ(brát)		*
b. ˈ(bráat)	*!	

Constraint ranking/strength

- No High insertion in /$\mu\mu$/: DEP-H $\gg$ HEAD-H

(10)

ʋuuk	DEP-H	HEAD-H
a. ☞ ˈ(ʋuuk)		*
b. ˈ(ʋúuk)	*!	

Gang effect

- Both HEAD-H and FTBIN are **individually** overridden by Faithfulness
- However, **joint** violation of HEAD-H and FTBIN is strictly prohibited

❖ GANGING-UP CUMULATIVITY: two weaker constraints jointly prevail against a stronger constraint (Jäger and Rosenbach 2006; Pater 2009b, 2016; Potts et al. 2010; Farris-Trimble 2008; Albright 2012; McPherson 2016; Ryan 2017; Breiss 2020; Breiss and Albright 2022)

Ganging up in BCMS

- DEP-$\mu \gg$ FTBIN
- HEAD-H&FTBIN $\gg$ DEP-μ

Ganging-up cumulativity in OT

- Not derivable in classical OT due to strict constraint ranking
- Modeled using LOCAL CONSTRAINT CONJUNCTION (Smolensky 1993, 2006; Itô and Mester 1998; Baković 2000)

(11) LOCAL CONSTRAINT CONJUNCTION $\mathbb{A}\&_{\mathscr{D}}\mathbb{B}$
Assess a mark for simultaneous violation of $\mathbb{A}$ and $\mathbb{B}$ within the designated local domain $\mathscr{D}$.

OT-LCC account of ML in BCMS

(12)

lɛd	Head-H&FtBin	Dep-μ	Head-H	FtBin
a. ˈ(lɛd)	*!		*	*
b. ☞ ˈ(lɛɛd)		*	*	

Gang effects in HG

- **Claim**: gang effects derivable in Harmonic Grammar (HG; Legendre, Miyata, and Smolensky 1990) **without LCC** (Potts et al. 2010; Pater 2009b, 2016; Farris-Trimble 2008)

(13)

	$\mathscr{H}$	$\mathbb{A}$ 2	$\mathbb{B}$ 1.5	$\mathbb{C}$ 1.5
a. ☞ abc	−2	−1		
b. def	−3		−1	−1

❖ **Advantage**: no need for LCC, which is prone to overgeneration (Kirchner 1996; Padgett 2002; McCarthy 2003; Pater 2009a, among others)

BCMS lengthening in standard HG

- A failed attempt of a standard HG analysis of vowel lengthening in toneless monosyllables in BCMS (14)

❖ DEP-μ outweighs FTBIN; cf. (9)

❖ HEAD-H indeterminate between candidates: both the faithful parse (a) and the intended winner (b) violate HEAD-H

⇒ The intended winner loses to the faithful parse regardless of HEAD-H's weight:

(14)

lɛd	$\mathscr{H}$	DEP-μ 2	HEAD-H n	FTBIN 1
a. 💣 ˈ(lɛd)	$-1+n$		-1	-1
b. ☹ ˈ(lɛɛd)	$-2+n$	-1	-1	

Asymmetric trade-off

- Asymmentric, i.e. one-to-many trade-off: stronger penalty traded against **multiple** weaker penalties:

(15) Asymmetric trade-off:

a. $w(\mathbb{A}) > w(\mathbb{B})$

b. $w(\mathbb{A}) > w(\mathbb{C})$

c. $w(\mathbb{B} + \mathbb{C}) > w(\mathbb{A})$

Pater 2009b, 2016; Jennifer L Smith 2022, 2018

Symmetric trade-off

- SYMMETRIC, i.e. ONE-TO-ONE TRADE-OFF: stronger penalty traded against **a single** weaker penalty (**BCMS**)

(16) SYMMETRIC TRADE-OFF

a. $w(\mathbb{A}) > w(\mathbb{B})$

b. $w(\mathbb{B}\&\mathbb{C}) > w(\mathbb{A}\&\mathbb{C})$

❖ BCMS: **symmetric trade-off**

(17) a. $w(\text{DEP-}\mu) > w(\text{FTBIN})$

b. $w(\text{HEAD-H+FTBIN}) > w(\text{HEAD-H+DEP-}\mu)$

Superadditivity

- Asymmetric trade-off → LINEARITY/ADDITIVITY → derivable in standard HG
- Symmetric trade-off → SUPERLINEARITY/SUPERADDITIVITY
  - Superadditivity is **not** derivable in standard HG
  - HG adopts weighted LCC to model superadditivity (Albright 2009; Green and Davis 2014; Shih 2016, 2017)

HG+LCC

(18)

lɛd	$\mathscr{H}$	DEP-μ 2	HD-H&FTBIN 1.5	HD-H n	FTBIN 1
a. ˈ(lɛd)	$-2.5+n$		-1	-1	-1
b. ☞ ˈ(lɛɛd)	$-2+n$	-1		-1	

- MaxEnt model comparison using the maxent.ot R package (Mayer, Tan, and Zuraw 2022)
- LCC significantly improves model fit: $\chi^2(1)$=555, p=.000
- Baseline loglik = -277
- Baseline+LCC loglik = -1.98227e-09

How powerful is too powerful?

- The use of LCC raises **typological concerns**:
  1. LCC's propensity to overgeneration (Kirchner 1996; Padgett 2002; McCarthy 2003; Pater 2009a)
  2. HG argued to be inherently **less restrictive than strict-ranking OT** (Prince and Smolensky 2004: 232–233; Prince and Smolensky 1997)
- Gap in the literature: OT+LCC vs HG+LCC as a theory of superadditivity

→ Typological survey:
  1. Is Head-H&FtBin overly powerful?
  2. Is HG+LCC more susceptible to overgeneration than OT+LCC

Methodology

- **Survey**: three constraint models:
  1. Baseline
  2. Baseline+LCC
  3. Baseline+TrochaicQuantity à la Zec (1999)
- Baseline:
  1. Head-H
  2. *Nonhead-H (de Lacy 2002);
  3. Nonfinality (moraic version) (Hyde 2003);
  4. FtBin;
  5. Stress-to-Weight;
  6. Weight-to-Stress;
  7. Max-H;
  8. Max-μ;
  9. Dep-H;
  10. Dep-μ;
  11. NoFlop

Zec 1999

(19) TrochaicQuantity
Assess a mark for every bimoraic foot whose head mora is higher-toned than the nonhead mora.

(20)

lɛd	TrochQuant	FtBin	Dep-μ
a. ˈ(lɛd)		$*$!	
b. ☞ ˈ(lɛɛd)			$*$

(21)

brát	TrochQuant	FtBin	Dep-μ
a. ☞ ˈ(brát)		$*$	
b. ˈ(bráat)	$*$!		$*$

Methodology

- Factorial typology of each model calculated using the OT-Help software (Staubs et al. 2010)
- Lists of grammars produced by OT-Help subjected to further scrutiny in R (R Core Team 2021)
- R code to aid the comparison of the models' predictions and the identification of pathological mappings

Supplementary materials

Results: monomoraic inputs

μ	$\acute{\mu}$	Grammar	Baseline	LCC	Zec (1999)
$ˈ(\mu)$	$ˈ(\acute{\mu})$	$\mathbb{F} \gg \mathbb{M}$	✓	✓	✓
$ˈ(\mu\mu)$	$ˈ(\acute{\mu}\mu)$	FtBin $\gg \mathbb{F}$	✓	✓	✓
$ˈ(\acute{\mu})$	$ˈ(\acute{\mu})$	Head-H $\gg \mathbb{F}$	✓	✓	✓
$ˈ(\acute{\mu}\mu)$	$ˈ(\acute{\mu}\mu)$	$\mathbb{M} \gg \mathbb{F}$	✓	✓	✓
$ˈ(\mu\mu)$	$ˈ(\acute{\mu})$	BCMS (1–3)	✗!	✓	✓
$ˈ(\mu\mu)$	$ˈ(\mu\mu)$	unattested	✗	✗	✓!

Results

⇒ LCC is **just as powerful as needed**:

1. it fixes the undergeneration issue of the baseline model;
2. it produces no pathological effects, unlike Zec (1999)'s model

- See Appendix 2 for other overgeneration effects of Zec (1999)

Empirical contribution

- Superadditivity in **categorical** processes; all previously reported superadditivity effects attested in **variable** patterns (Shih 2017; B. W. Smith and Pater 2020; Breiss and Albright 2022; Kim 2022)
- **Tone-sensitive prosodic minimality**: tone not only determines the locus of stress (Goldsmith 1987; de Lacy 2002; Gordon 2023), but also dictates the minimal acceptable shape of prosodic constituents

Theoretical contribution

- **H**G+LCC produces desirable factorial typology:
  1. more powerful than standard HG (no superadditivity);
  2. more restrictive (!) than Zec (1999)'s model
- HG+LCC is **not** inherently more prone to overgeneration than OT+LCC

❖ Conjunction of two functionally related prosodic constraints ⇒ constraints on which constraint types can be conjoined

References I

Albright, Adam (2009). "Cumulative violations and complexity thresholds: evidence from Lakhota". Paper presented at the 17th Manchester Phonology Meeting.

— (2012). "Additive markedness interactions in phonology". Unpublished manuscript, MIT.

Baković, Eric (2000). "Harmony, dominance and control". PhD thesis. New Brunswick, NJ: Rutgers University.

Breiss, Canaan (2020). "Constraint cumulativity in phonotactics: Evidence from artificial grammar learning studies". *Phonology* 37.4, pp. 551–576.

Breiss, Canaan and Adam Albright (2022). "Cumulative markedness effects and (non-) linearity in phonotactics". *Glossa: a journal of general linguistics* 7.1.

de Lacy, Paul (2002). "The interaction of tone and stress in Optimality Theory". *Phonology* 19.1, pp. 1–32. DOI: https://doi.org/10.1017/S0952675702004220.

Farris-Trimble, Ashley W (2008). "Cumulative faithfulness effects in phonology". PhD thesis. Indiana University.

Goldsmith, John (1987). "Tone and accent, and getting the two together". In: *Annual Meeting of the Berkeley Linguistics Society*. Vol. 13, pp. 88–104.

Gordon, Matthew (2023). "The Phonetic Basis for Tone-Stress Interactions: A Cross-Linguistic Study". In: *Syllable, Stress, and Sign*. Ed. by Jeroen van de Weijer. Walter de Gruyter GmbH & Co KG, pp. 171–190.

Green, Christopher R and Stuart Davis (2014). "Superadditivity and limitations on syllable complexity in Bambara words". In: *Perspectives on phonological theory and development, in honor of Daniel A. Dinnsen*. Ed. by Ashley W. Farris-Trimble and Jessica A. Barlow, pp. 223–247.

Hyde, Brett (2003). "Nonfinality". Unpublished manuscript, Washington University. URL: http://roa.rutgers.edu/article/view/643.

Itô, Junko and Armin Mester (1998). "Markedness and word structure: OCP effects in Japanese". Manuscript, University of California, Santa Cruz.

Jäger, Gerhard and Anette Rosenbach (2006). "The winner takes it all—almost: Cumulativity in grammatical variation". *Linguistics* 44, pp. 937–971.

References II

Kim, Seoyoung (2022). "A MaxEnt learner for super-additive counting cumulativity". *Glossa: a journal of general linguistics* 7.1.

Kirchner, Robert (1996). "Synchronic chain shifts in Optimality Theory". *Linguistic Inquiry* 27.2, pp. 341–350.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky (1990). "Harmonic Grammar—A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations". In: *Proceedings of the twelfth annual conference of the Cognitive Science Society*. Citeseer, pp. 884–891.

Mayer, C., A. Tan, and K. Zuraw (2022). *maxent.ot: A package for doing Maximum Entropy Optimality Theory in R (Version 0.1.0)*. Computer software. URL: https://doi.org/10.5281/zenodo.7246367.

McCarthy, John J (2003). "Comparative markedness". *Theoretical Linguistics* 29, pp. 1–51.

McCarthy, John J and Alan Prince (1995). "Faithfulness and reduplicative identity". In: *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. URL: http://works.bepress.com/john_j_mccarthy/44/.

McPherson, Laura (2016). "Cumulativity and ganging in the tonology of Awa suffixes". *Language* 92.1, e38–e66.

Padgett, Jaye (2002). "Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory". Unpublished manuscript, University of California Santa Cruz.

Pater, Joe (2009a). "Paul Smolensky and Géraldine Legendre (2006). The harmonic mind: from neural computation to optimality-theoretic grammar. Cambridge, Mass.: MIT Press. Vol. 1: Cognitive architecture. Pp. xxiv+ 563. Vol. 2: Linguistic and philosophical implications. Pp. xxiv+ 611.". *Phonology* 26.1, pp. 217–226.

— (2009b). "Weighted constraints in generative linguistics". *Cognitive science* 33.6, pp. 999–1035.

— (2016). "Universal grammar with weighted constraints". In: *Harmonic grammar and harmonic serialism*. Ed. by Joe Pater and John McCarthy. London: Equinox, pp. 1–46.

Potts, Christopher et al. (2010). "Harmonic Grammar with linear programming: from linear systems to linguistic typology". *Phonology* 27.1, pp. 77–117.

References III

Prince, Alan (1990). "Quantitative consequences of rhythmic organization". In: *Papers from the 26th Regional Meeting of the Chicago Linguistic Society: Volume 2: The Parasession*. Ed. by Michael Ziolkowski, Manuela Noske, and Karen Deaton. Chicago Linguistic Society, pp. 355–398.

Prince, Alan and Paul Smolensky (1997). "Optimality: From neural networks to universal grammar". *Science* 275.5306, pp. 1604–1610.

— (2004). *Optimality Theory: Constraint interaction in generative grammar*. John Wiley & Sons.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Ryan, Kevin M. (2017). "Attenuated spreading in Sanskrit retroflex harmony". *Linguistic inquiry* 48.2, pp. 299–340.

Shih, Stephanie S (2016). "Super additive similarity in Dioula tone harmony". In: *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, pp. 361–370.

— (2017). "Constraint conjunction in weighted probabilistic grammar". *Phonology* 34.2, pp. 243–268.

Smith, Brian W and Joe Pater (2020). "French schwa and gradient cumulativity". *Glossa: a journal of general linguistics* 5.1.

Smith, Jennifer L (2022). "Some Formal Implications of Deletion Saltation". *Linguistic Inquiry* 53.4, pp. 852–864.

— (2018). "Feature change is not like deletion: Saltation in Harmonic Grammar". Talk at LSA 2018. URL: https://users.castle.unc.edu/~jlsmith/home/pdf/smith2018_SaltationInHG_LSA-slides.pdf.

Smolensky, Paul (1993). "Harmony, markedness, and phonological activity". Paper presented at Rutgers Optimality Workshop.

References IV

Smolensky, Paul (2006). “Optimality in phonology II: harmonic completeness, local constraint conjunction, and feature domain markedness”. In: *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar. Vol. 2: Linguistic and pholosophical implications.* Ed. by Paul Smolensky and Géraldine Legendre. Massachusetts Institute of Technology Press, pp. 27–160.

Staubs, Robert et al. (2010). “OT-Help 2.0”. Software package, University of Massachusetts Amherst. URL: https://people.umass.edu/othelp/.

Yip, Moira (2001). “The complex interaction of tone and prominence”. In: *NELS*. Vol. 31. 2, pp. 531–545.

Zec, Draga (1999). “Footed tones and tonal feet: rhythmic constituency in a pitch-accent language”. *Phonology* 16.2, pp. 225–264. DOI: https://doi.org/10.1017/S0952675799003759.

Acknowledgments

Thanks to:

- Kevin Ryan and Aleksei Nazarov
- Audiences at Harvard PhonLab meetings (Fall 2022, Fall 2023), ECO-5 (UConn), CLS 59, FASL 32 (Indiana University Bloomington)

Tonal Flop

Some Old Štokavian dialects of BCMS (Ivić 1958): e.g., Trstenik (Jović 1968), Resava (Peco and Milanović 1968)

(22) /ʋɔd-á/ ‘water-NOM.SG’

a.	(ˈʋɔ́.da)	Tonal Flop	**light penult**
b.	ʋɔ.(ˈdá=jɛ)	‘water=is’	Enclisis

(23) /ruuk-á/ ‘arm-NOM.SG’

a.	ˈ(ruu).ká	**no** Tonal Flop	**heavy penult**
b.	ruu.(ˈká=jɛ)	‘arm=is’	Enclisis

Tonal Flop

Faithful parse	Ruled out because	Penalized by
ʋɔ.ˈ(dá)	stressed final light	NONFINALITY
(ˈʋɔ.dá)	High-toned nonhead mora	*NONHEAD-H
ˈ(ʋɔ).dá	**toneless monomoraic foot**	HEAD-H&FTBIN

TABLE 1. Failed faithful parses

- Highs do not shift to the head mora of a bimoraic foot (23a)

⇒ NOFLOP-H ≫ HEAD-H

Tonal Flop

ʋɔdá		$\mathscr{H}$	NONFIN 4	*NONHD-H 4	DEP-μ 4	HD-H&FTB 2	NOFLOP 2	HD-H 1.5	FTBIN 1
a.	ʋɔˈ(dá)	−5	−1						−1
b.	(ˈʋɔdá)	−5.5		−1				−1	
c.	ˈ(ʋɔ)dá	−4.5				−1		−1	−1
d. ☞	(ˈʋɔ́da)	−2					−1		
e.	ˈ(ʋɔɔ)dá	−5.5			−1			−1	

Penultimate Lengthening

Old Štokavian dialects that do not feature Tonal Flop (e.g., Gallipoli Serbian (Ivić 1957/1994)

(24) /ʋɔd-á/ ‘water-NOM.SG’

a.	ˈ(ʋɔɔ).dá	Lengthening	**toneless light**
b.	ʋɔ.(ˈdá=jɛ)	‘water=is’	Enclisis

(25) No lengthening in High-toned monomoraic feet

/kɔ́maata/ ‘piece.GEN.SG’ ˈ(kɔ́).maa.ta

Penultimate Lengthening

Faithful parse	Ruled out because	Penalized by
ʊɔ.ˈ(dá)	stressed final light	NONFINALITY
(ˈʊɔ.dá)	High-toned nonhead mora	*NONHEAD-H
ˈ(ʊɔ).dá	**toneless monomoraic foot**	HEAD-H&FTBIN

TABLE 2. Failed faithful parses

- No lengthening by default; *not all* monomoraic feet (25) are eliminated

⇒ DEP-$\mu \gg$ FTBIN

Penultimate Lengthening

	υɔdá	$\mathscr{H}$	NonFin 4	*NonHd-H 4	NoFlop 4	Hd-H&FtB 2	Dep-μ 2	Hd-H 1.5	FtB 1
a.		υɔˈ(dá)	−5	−1						−1
b.		(ˈυɔdá)	−5.5		−1				−1	
c.		ˈ(υɔ)dá	−4.5				−1		−1	−1
d.		(ˈυɔ́da)	−4			−1				
e.	☞	ˈ(υɔɔ)dá	−3.5					−1	−1	

Zec (1999)'s predictions

- Zec 1999: Vowel lengthening in High-toned monosyllables in BCMS (1) is blocked by the TROCHQUANT constraint, which penalizes bimoraic trochees whose head mora has a higher tone than the nonhead mora

⇒ Penalize the combination of two unmarked structures

- TROCHQUANT **overgenerates** (in what follows, I report **only a subset** of the pathological effects observed in the typological survey)

Pathology #1

Shortening effect: TROCHQUANT $\gg$ MAX-μ

1. Stressed **High-toned** heavies become light:

$\acute{\mu}\mu$		TRQ	MAX-H	MAX-μ	FTBIN
a.	ˈ($\acute{\mu}\mu$)	*!			
b.	ˈ($\mu\mu$)		*!		
c. ☞	ˈ($\acute{\mu}$)			*	*

Pathology #1

2. Stressed **toneless** heavies surface faithfully:

$\mu\mu$			TRQ	MAX-μ	FTBIN
a.	☞	$^{\text{ˈ}}(\mu\mu)$			
b.		$^{\text{ˈ}}(\mu)$		*!	*!

Pathology #2

High tone deletion: TrochQuant $\gg$ Max-H

1. High tone deletes in the head mora of a **bimoraic** trochee:

$\acute{\mu}\mu$		TrochQuant	Max-μ	Max-H
a.	$'(\acute{\mu}\mu)$	*!		
b.	$'(\acute{\mu})$		*!	
c. ☞	$'(\mu\mu)$			*

Pathology #2

2. High tone does **not** delete in the head position of a **monomoraic** foot

	$\acute{\mu}$		TRQ	MAX-H
a.	☞	$\text{ˈ}(\acute{\mu})$		
b.		$\text{ˈ}(\mu)$		*!