Noisy HG Models of Eastern Andalusian Harmony*

Aaron Kaplan
University of Utah
a.kaplan@utah.edu

OCP 16
Jan. 18, 2019

1 Introduction

- Noisy Harmonic Grammar (NHG): depending on when and where noise is added, different patterns of variation are possible (Hayes 2017).
- ATR harmony in Eastern Andalusian as a test of these possibilities: variation is constrained by categorical requirements; depending on the constraint set, some attested forms are harmonically bounded.
- The best model perturbs constraint weights at the outset of the evaluation.
  - This version of NHG cannot produce harmonically bounded forms and therefore requires constraints under which attested forms are not harmonically bounded.

2 Eastern Andalusian Harmony

- /s/-aspiration: word-final /s/ deletes, triggering laxing of adjacent vowel. These lax vowels trigger variable harmony on preceding vowels.
- The stressed vowel always harmonizes (data from Jiménez & Lloret 2007, Lloret & Jiménez 2009):

(1)

a.	*tesis*	tɛ́sɪ	'thesis'
b.	*tienes*	tjɛ́nɛ	'you have'
c.	*nenes*	nɛ́nɛ	'babies'
d.	*monos*	mɔ́nɔ	'monkeys'
e.	*lejos*	lɛ́hɔ	'far'
f.	*pesos*	pɛ́sɔ	'weights'
g.	*bocas*	bɔ́kæ̝	'mouths'

*I am grateful to participants in the Analyzing Typological Structure workshop at Stanford University (Sept. 22, 2018) for feedback on this work, and thanks especially to Abby Kaplan for assistance with R.

- Other post-tonic vowels optionally harmonize as a group:

(2)

a.	*treboles*	tréβolɛ ∼ tréβɔlɛ	‘clovers’
b.	*cómetelos*	kómetelɔ ∼ kómɛtɛlɔ	‘eat them (for you)!’
		*kómɛtelɔ, *kómetɛlɔ	

- Pretonic vowels optionally harmonize as a group, but only with post-tonic harmony:

(3)

a.	*momentos*	moméntɔ ∼ mɔméntɔ	‘instants’
b.	*reloj*	relɔ́ ∼ rɛlɔ́	‘watch’
c.	*relojes*	relɔ́hɛ ∼ rɛlɔ́hɛ	‘watches’
d.	*monederos*	moneðéɾɔ ∼ mɔnɛðéɾɔ	‘purses’
		*mɔneðéɾɔ, *monɛðéɾɔ	
e.	*cojines*	kohínɛ ∼ kɔhínɛ	‘pillows’
f.	*cotillones*	kotiʒɔ́nɛ ∼ kɔtiʒɔ́nɛ	‘cotillions’
g.	*recógelos*	rekɔ́helɔ ∼ rekɔ́hɛlɔ ∼ rɛkɔ́hɛlɔ	‘pick them’
		*rɛkɔ́helɔ	

- High vowels lax word finally but do not undergo harmony:

(4)

a.	*crisis*	krísɪ	‘crisis’
b.	*muchos*	múʃɔ	‘many’
c.	*mios*	míɔ	‘mine (pl.)’

- Positional licensing (PL): [–ATR] must appear in the stressed syllable or in every syllable (Jiménez & Lloret 2007, Lloret 2018, Lloret & Jiménez 2009, Walker 2011; analyses below are based on this work).

- Pairing different versions of PL with particular implementations of NHG gives a range of models of Eastern Andalusian.
  - Constraints: negative and positive versions of PL (Kaplan 2018).
  - NHG: 8 implementations from Hayes (2017).

- The best model: Hayes’s (2017) “classic NHG” with positive PL
  - Classic NHG: no harmonically bounded outputs. Therefore, no licit output can be harmonically bounded (the case under positive but not negative PL).
  - Other versions of NHG produce harmonically bounded forms, but cannot distinguish “good” ones from “bad” ones.

3 Positional Licensing Analyses

3.1 Candidates of Interest

(5)

Input	Candidate	Attested?	Neg. PL	Pos. PL
a. /monedéros/	moneðéro			
'purses'	moneðérɔ			
	moneðɛ́rɔ	✓		
	mɔneðɛ́rɔ		Bounded	Bounded
	monɛðɛ́rɔ		Bounded	Bounded
	mɔnɛðɛ́rɔ	✓		
b. /kómetelos/	kómetelo			
'eat them	kómetelɔ			
(for you)!'	kɔ́metelɔ	✓	Bounded	
	kɔ́mɛtelɔ		Bounded	Bounded
	kɔ́metɛlɔ		Bounded	Bounded
	kɔ́mɛtɛlɔ	✓		
c. /rekógelos/	rekóhelo			
'pick them'	rekóhelɔ			
	rekɔ́helɔ	✓	Bounded	
	rekɔ́hɛlɔ	✓		
	rɛkɔ́helɔ		Bounded	Bounded
	rɛkɔ́hɛlɔ	✓		
d. /krísis/	krísi			
'crisis'	krísɪ	✓		
	krísɪ			

- Positive PL: no attested candidate is harmonically bounded.
- Negative PL: two attested candidates are harmonically bounded: *kɔ́metelɔ, rekɔ́helɔ*.
- Both: some unattested candidates are harmonically bounded; other are not.
- NHG with negative PL must produce *kɔ́metelɔ, rekɔ́helɔ* without producing other harmonically bounded forms.

3.2 Negative PL

- To avoid pathologies in HG, PL must be gradient: Negative Gradient PL (NG-PL; Kaplan 2018):

(6) LICENSE([–ATR], $\acute{\sigma}$): assign -1 for each [–ATR] that does not coincide with $\acute{\sigma}$ and -1 for each syllable that intervenes between [–ATR] and the nearest $\acute{\sigma}$.

- This accounts for harmony up to the licensor.

- Pretonic harmony: Maximal Licensing (Walker 2011) requires [–ATR] to appear in every syllable.
- IDENT(ATR) disfavors harmony.
- These constraints produce post-tonic and pretonic harmony, but forms with no post-tonic harmony are harmonically bounded.
- LICENSE penalizes unharmonized post-tonic vowels in *kɔ́metelɔ*, *rekɔ́helɔ* to avoid pathologies (Kaplan 2018).
- ☞ = attested; × = harmonically bounded

(7) a.

/monedéros/	LICENSE	MAXLIC	IDENT	*Comments*
a. moneðéɾɔ	−1	−3	−1	
☞ b. moneðéɾɔ		−2	−2	
☞ c. mɔnɛðéɾɔ			−4	
× d. mɔneðéɾɔ		−1	−3	collectively bounded[1] by (b) & (c)
× e. monɛðéɾɔ		−1	−3	collectively bounded by (b) & (c)

b.

/kómetelos/	LICENSE	MAXLIC	IDENT	*Comments*
a. kómetelɔ	−3	−3	−1	
×☞ b. kɔ́metelɔ	−2	−2	−2	collectively bounded by (a) & (c)
☞ c. kɔ́mɛtɛlɔ			−4	
× d. kɔ́mɛtelɔ	−1	−1	−3	collectively bounded by (a) & (c)
× e. kɔ́metɛlɔ	−1	−1	−3	collectively bounded by (a) & (c)

c.

/rekóhelos/	LICENSE	MAXLIC	IDENT	*Comments*
a. rekóhelɔ	−2	−3	−1	
×☞ b. rekɔ́helɔ	−1	−2	−2	collectively bounded by (a) & (c)
☞ c. rekɔ́hɛlɔ		−1	−3	
☞ d. rɛkɔ́hɛlɔ			−4	
× e. rɛkɔ́helɔ	−1	−1	−3	bounded by (c)

[1]Collective harmonic bounding: Samek-Lodovici & Prince (1999)

- High vowels: *[+hi, –ATR] prevents harmony, MAX(–ATR) forces laxing word-finally.

(8)

/kɾ́isis/	*[+hi, –ATR]	MAX(–ATR)	LICENSE	MAXLIC	IDENT
a. kɾ́isi		−1			
☞ b. kɾ́isɪ	−1		−1	−1	−1
c. kɾísɪ	−2				−2

- What to do about the harmonically bounded attested forms?
  - Nothing: let NHG deal with them.
  - Revise PL: Positive Gradient PL (PG-PL; Kaplan 2018)

3.3 Positive PL

(9) LICENSE([–ATR], $\acute{\sigma}$): assign +1 for each [–ATR] that coincides with $\acute{\sigma}$ and +1 for each additional syllable that [–ATR] appears in.

- This subsumes MAXLIC; we need IDENT(ATR)-pretonic to block pretonic harmony.
- All attested forms are now possible winners.

(10) a.

/monedéros/	LICENSE	IDENT-pretonic	IDENT
a. moneðéɾɔ			−1
☞ b. moneðéɾɔ	+2		−2
☞ c. mɔnɛðéɾɔ	+4	−2	−4
× d. mɔneðéɾɔ	+3	−1	−3
× e. monɛðéɾɔ	+3	−1	−3

b.

/kómetelos/	LICENSE	IDENT-pretonic	IDENT
a. kómetelɔ			−1
☞ b. kɔ́metelɔ	+2		−2
☞ c. kɔ́mɛtɛlɔ	+4		−4
× d. kɔ́mɛtelɔ	+3		−3
× e. kɔ́metɛlɔ	+3		−3

c.

/rekóhelos/	License	Ident-pretonic	Ident
a. rekóhelɔ			−1
☞ b. rekɔ́helɔ	+2		−2
☞ c. rekɔ́hɛlɔ	+3		−3
☞ d. rɛkɔ́hɛlɔ	+4	−1	−4
× e. rɛkɔ́helɔ	+3	−1	−3

d.

/krísis/	*[+hi, –ATR]	Max(–ATR)	License	Ident-pretonic	Ident
a. krísi		−1			
☞ b. krísɪ	−1				−1
☞ c. krísɪ	−2		+2		−2

(11) Core weighting requirements:

a. Harmony on $\acute{\sigma}$ only: $2w(\text{License}) > w(\text{Ident}) > w(\text{License})$
b. Full post-tonic harmony: $w(\text{Ident}) + w(\text{Ident-pre}) > w(\text{License}) > w(\text{Ident})$
c. Maximal harmony: $w(\text{License}) > w(\text{Ident}) + w(\text{Ident-pretonic})$
d. High vowels: $w(\text{Max}(\text{–ATR})) > w(*[+\text{hi}, \text{–ATR}]) + w(\text{Ident}) > 2w(\text{License})$

- Summary: 2 ways to account for Eastern Andalusian:

1. NG-PL: NHG responsible for variation and relieving harmonic bounding.
2. PG-PL: NHG responsible for variation only.

4 Simulations

- Monte Carlo simulations following Hayes (2017) using OTSoft (Hayes et al. 2013): 8 variants of NHG; NG-PL and PG-PL.

1. Noise at the constraint level
   (a) Classic NHG: Noise added before multiplication of penalties by weights: $penalty * (weight + noise)$
   (b) Noise added after multiplication of penalties by weights, no noise allowed if penalty = 0: $(penalty * weight) + noise$
   (c) Noise added after multiplication of penalties by weights, noise allowed if penalty = 0: $(penalty * weight) + noise$
2. Noise at the cell level
   (a) Noise added before multiplication of penalties by weights: $penalty * (weight + noise)$
   (b) Noise added after multiplication of penalties by weights, no noise allowed if penalty = 0: $(penalty * weight) + noise$

(c) Noise added after multiplication of penalties by weights, noise allowed if penalty = 0: $(penalty * weight) + noise$

3. Noise at the candidate level
4. Maximum Entropy (Goldwater & Johnson 2003)

- 100,000 trials per simulation. Negative constraint weights were disallowed.
- Most successful arrangement: Hayes's classic NHG (variety 1a) with PG-PL:

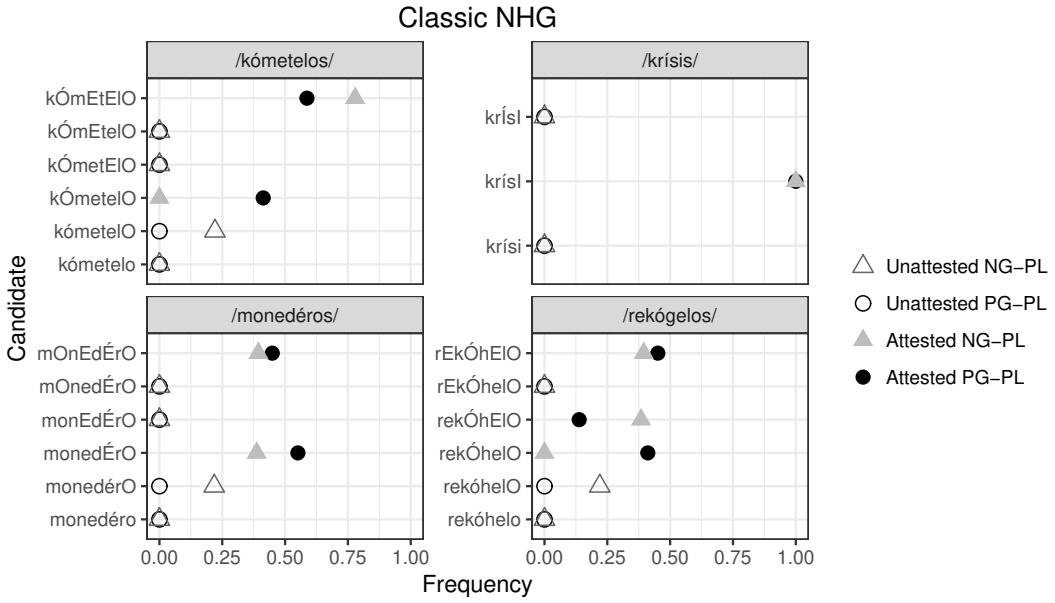

Figure 1: Results of simulations under variety 1a

- In particular simulation shown here, all and only attested forms produced. Not a minor accomplishment: some illicit forms are not harmonically bounded.
- Subsequent simulations: unattested forms produced rarely. Worst result: *krísi* produced 38 times out of 100,000 trials. 2 other illicit forms produced: *kɔ́mɛtelɔ, monɛðérɔ*
- The same simulation with NG-PL is less successful: classic NHG effectively does not produce harmonically bounded candidates.[2]
  - Attested [kɔ́metelɔ], [rekɔ́helɔ] cannot be produced.
  - Unattested [moneðérɔ], [kómetelɔ], [rekóhelɔ] appear at a $\sim 22\%$ rate.

[2]With only positive constraint weights, a harmonically bounded candidate is selected under classic NHG only when it ties with a rival (Hayes 2017). Ties occurred very rarely in my simulations (for the PG-PL simulation in Figure 1: 125 ties in 66,565,284 chances), so I take it to be a reasonable approximation to say that classic NHG does not produce harmonically bounded candidates. Indeed, in none of my simulations with classic NHG did a harmonically bounded candidate win.

- Classic NHG succeeds only when no attested form is harmonically bounded. Under those conditions, it performs very well on Eastern Andalusian.

4.1 Constraint-Level Noise

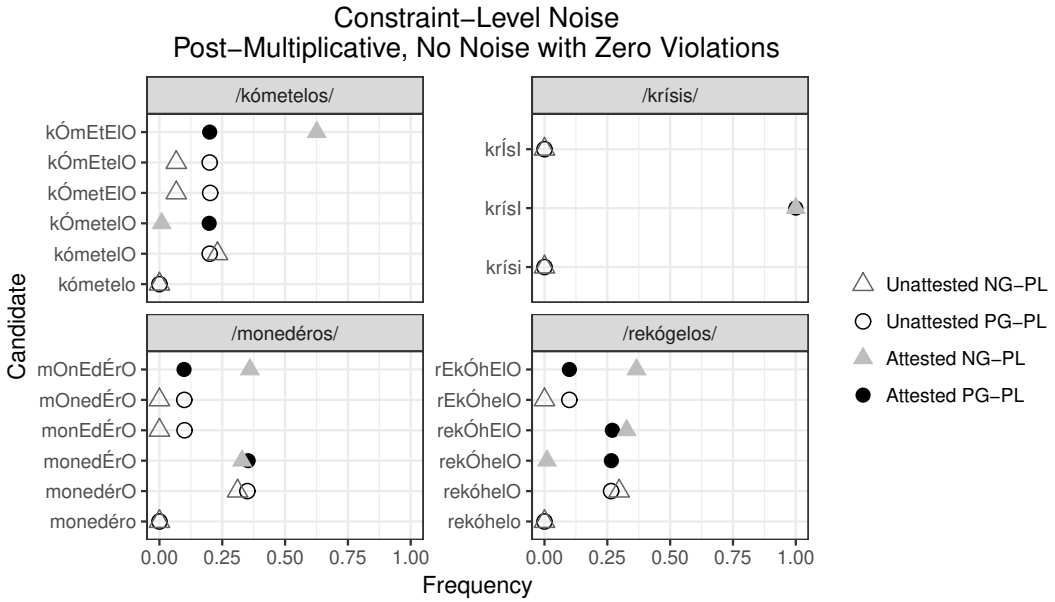

Figure 2: Results of simulations under variety 1b

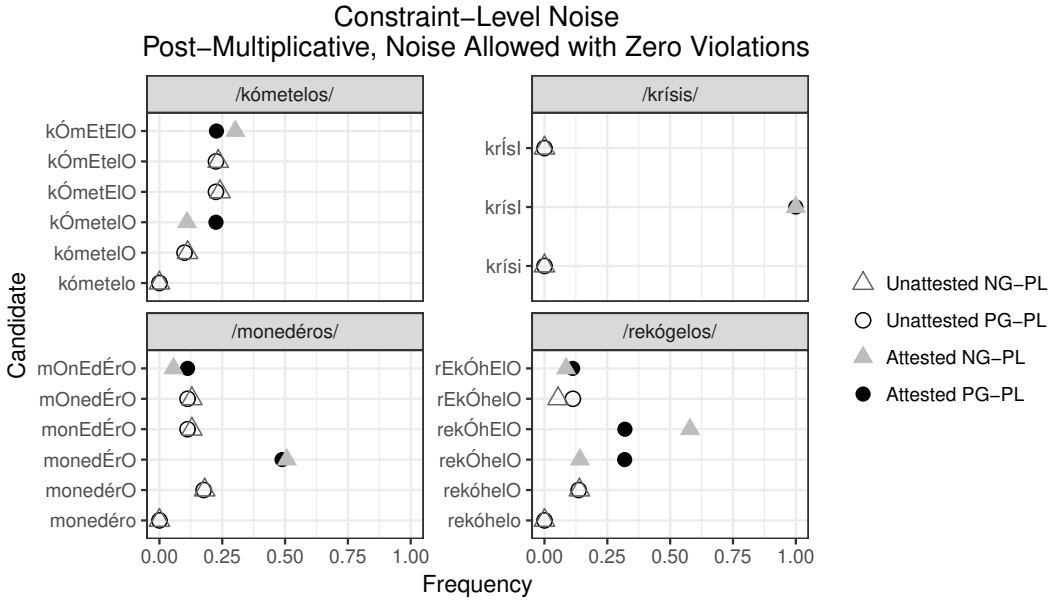

Figure 3: Results of simulations under variety 1c

4.2 Cell-Level Noise

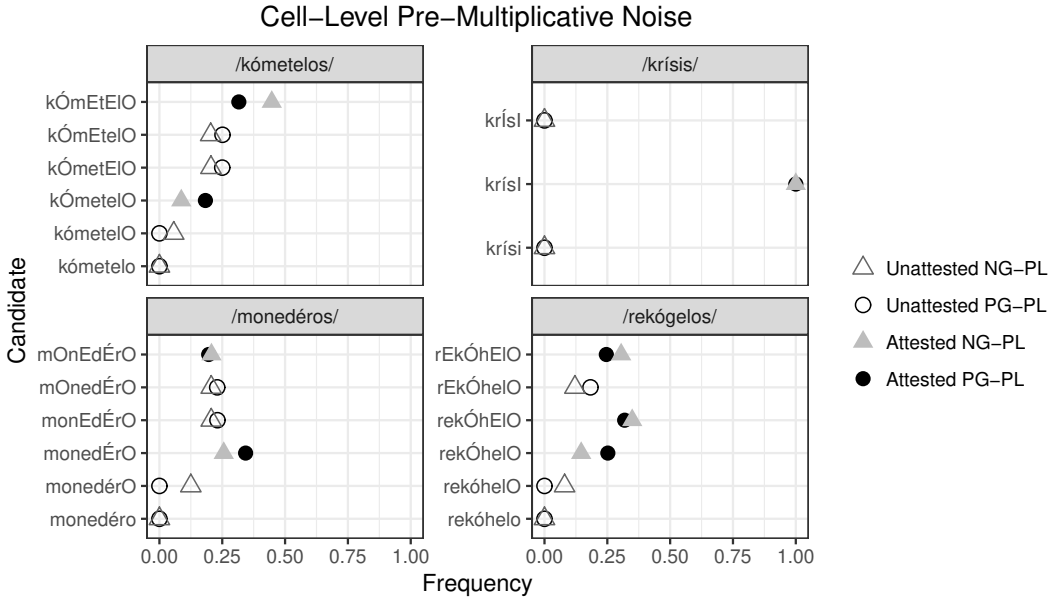

Figure 4: Results of simulations under variety 2a

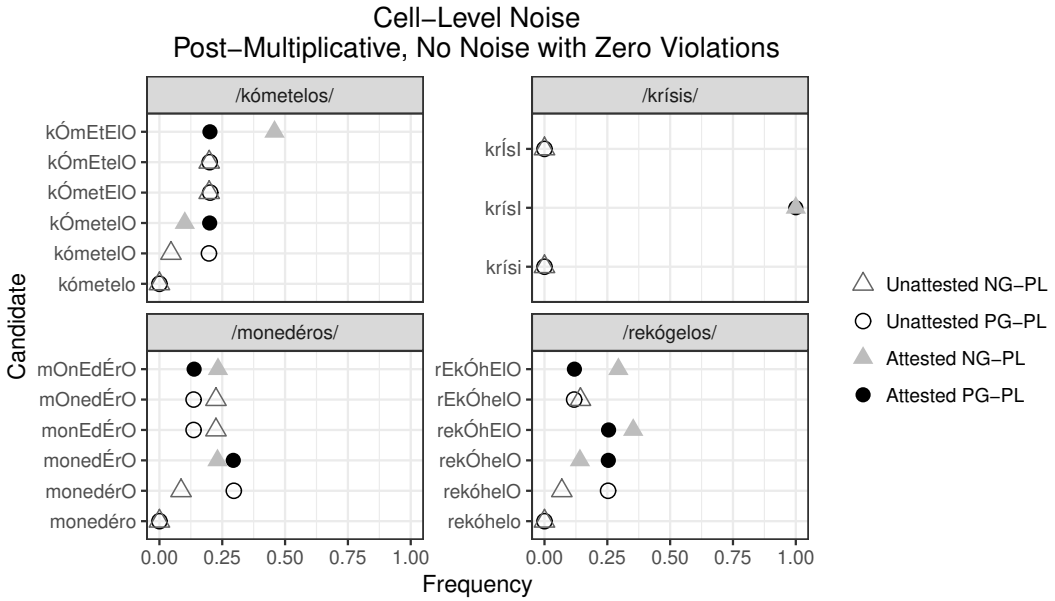

Figure 5: Results of simulations under variety 2b

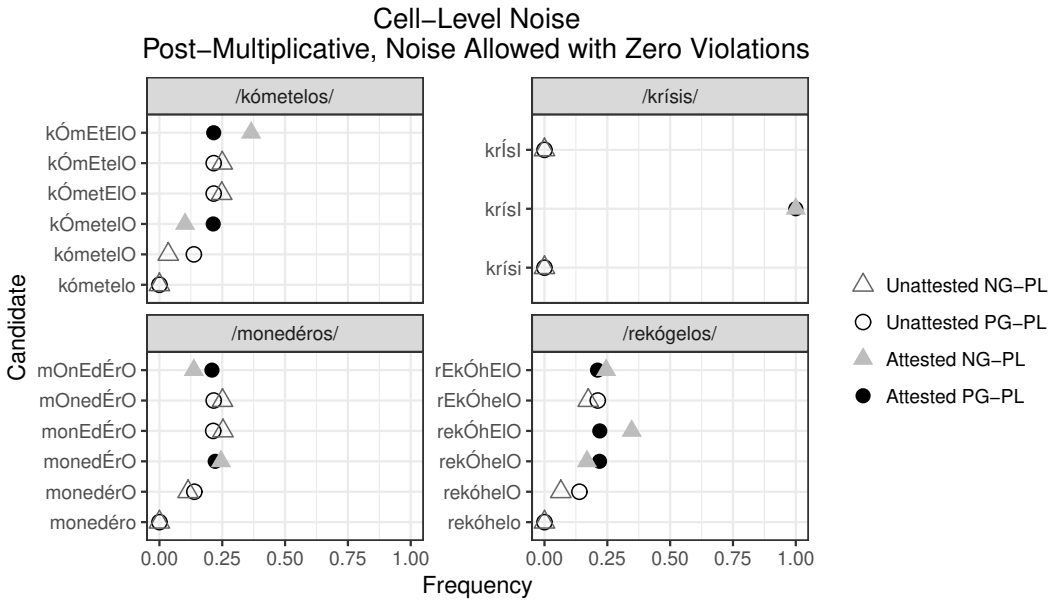

Figure 6: Results of simulations under variety 2c

4.3 Candidate-Level Noise

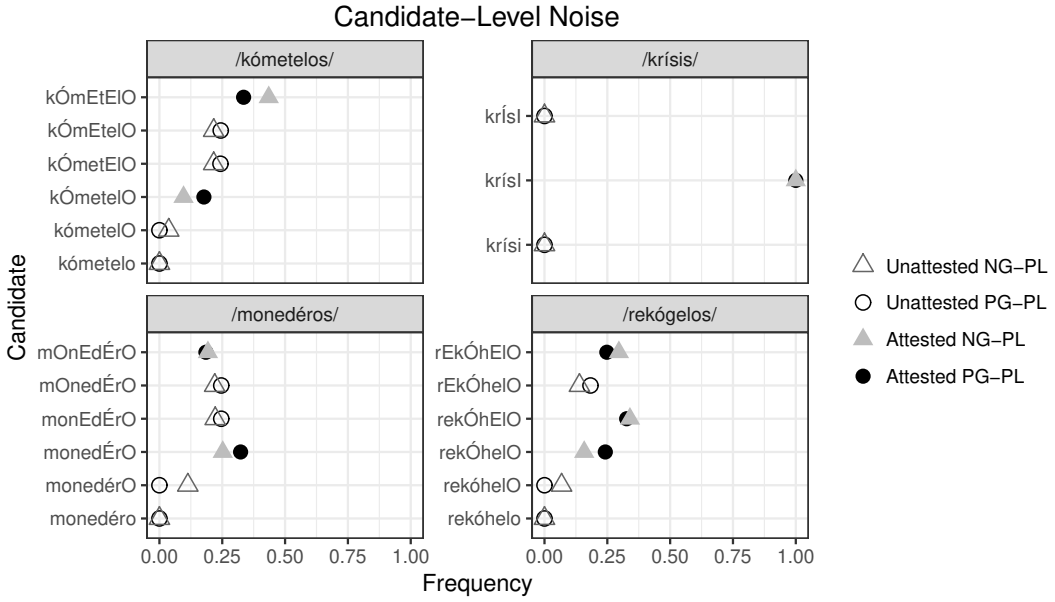

Figure 7: Results of simulations under variety 3

4.4 MaxEnt

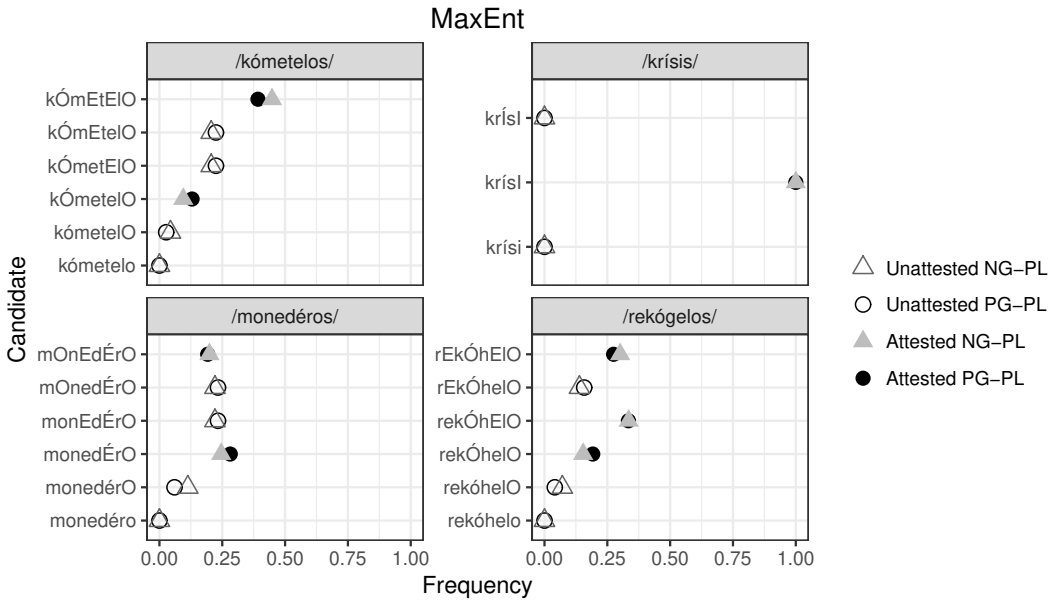

Figure 8: Results of simulations under MaxEnt

5 Discussion

- /kŕisis/: no variation here, so weights approximating "MAX(–ATR) ≫ *[+hi, –ATR] ≫ everything else" can be established.
- For this reason, forms with no lax vowels (e.g. *moneðéro*) never win.
- The nature of Eastern Andalusian's optionality is tailor-made for classic NHG:
  - Many unattested candidates are harmonically bounded and therefore inaccessible to classic NHG.
  - Remaining unattested forms: no lax vowels (e.g. *moneðéro*), ruled out by high-weighted MAX(–ATR); no harmony (*moneðéɾɔ*), ruled out if IDENT does not outweigh LICENSE by too much.
  - The weights found under this simulation reflect these criteria:

(12)

46.000	MAX(–ATR)
27.000	*[+hi, –ATR]
11.655	LICENSE
11.345	IDENT(ATR)
0.251	IDENT(ATR)-pretonic

- Other implementations of NHG make it easier to subvert these arrangements: harmonically bounded candidates can win, or crucial weighting relationships can be reversed (e.g. by adding noise unequally to candidates).

- NHG cannot relieve the harmonic-bounding problem on its own: opening the door to one bounded candidate opens the door to others.
- Better to let the constraints identify viable candidates that NHG can choose from.

6 Conclusion

- These results provide support for classic NHG and positive constraints.
- The differences between versions of NHG can be subtle, but it is possible to distinguish them empirically.
- Small changes make a big difference.

References

Goldwater, Sharon & Mark Johnson (2003) Learning OT Constraint Rankings Using a Maximum Entropy Model. In *Proceedings of the Workshop on Variation within Optimality Theory*, 113–122, Stockholm University.

Hayes, Bruce (2017) Varieties of Noisy HG. In *Proceedings of AMP 2016*, Karen Jesney, Charlie O'Hara, Caitlin Smith, & Rachel Walker, eds., Washington, DC: Linguistic Society of America.

Hayes, Bruce, Bruce Tesar, & Kie Zuraw (2013) OTSoft 2.5. software package, http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Jiménez, Jesús & Maria-Rosa Lloret (2007) Andalusian Vowel Harmony: Weak Triggers and Perceptibility. paper presented at the 4th Old World Conference in Phonology, Workshop on Harmony in the Languages of the Mediterranean, Rhodes, January 18-21, 2007.

Kaplan, Aaron (2018) Positional Licensing, Asymmetric Trade-Offs, and Gradient Constraints in Harmonic Grammar. *Phonology* **35**: 247–286.

Lloret, Maria-Rosa (2018) Andalusian Vowel Harmony at the Phonology-Morphology Interface. Talk presented at the 2015 Old World Conference on Phonology, London, January 12-14.

Lloret, Maria-Rosa & Jesús Jiménez (2009) Un Análisis *Óptimo* de la Armonía Vocálica del Andaluz. *Verba* **36**: 293–325.

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

Walker, Rachel (2011) *Vowel Patterns in Language*. New York: Cambridge University Press.