Gradual Optionality in Noisy HG

Aaron Kaplan

January 6, 2021

- Noisy Harmonic Grammar: probabilistic implementations of Harmonic Grammar (Jesney 2007; Hayes 2017; Flemming 2017; Zuraw & Hayes 2017).
- **Question 1:** How do its properties change when implemented serially?
- **Question 2:** Is it compatible with gradual approaches to deletion and feature change (McCarthy 2008)?

Introduction

- Only one version of NHG supports an analysis of optionality in Eastern Andalusian harmony in a parallel framework (Kaplan 2018a; Kaplan 2019).
  - The harmony-driving constraint in this analysis requires serialism.
- Harmony can be implemented gradually.
  - Does noise interfere with the necessary sequence of steps?

Introduction

Two test cases:

- **Eastern Andalusian harmony (Jiménez & Lloret 2007; Lloret & Jiménez 2009; Lloret 2018):** the same implementation of NHG that succeeds in parallel is also the only one that succeeds in serialism.
- **Hiatus resolution in Persian (Ariyaee & Jurgec 2020):** NHG successfully produces gradual vowel deletion; modeling output frequencies using serial NHG requires revisions to particular constraints.

Serial NHG closely resembles parallel NHG, but we may need to rethink our constraints.

Variable Harmony in Eastern Andalusian

- /s/-aspiration (= deletion) causes laxing of word final vowel, which triggers [–ATR] harmony on the stressed syllable:

tesis	ˈtɛsɪ	‘thesis’	*nenes*	ˈnɛnɛ	‘babies’
tienes	ˈtjɛnɛ	‘you have’	*pesos*	ˈpɛsɔ	‘weights’
monos	ˈmɔnɔ	‘monkeys’	*lejos*	ˈlɛhɔ	‘far’
bocas	ˈɔkæ̞	‘mouths’	*asas*	ˈa̞sæ̞	‘handles’

- Harmony on other vowels is optional. . .

Variable Harmony in Eastern Andalusian

- Nonfinal post-tonic vowels optionally harmonize in lockstep:

treboles	ˈtrɛβolɛ ∼ ˈtrɛβɔlɛ	'clovers'
cómetelos	ˈkɔmetelɔ ∼ ˈkɔmɛtɛlɔ	'eat them (for you)!'
	*ˈkɔmɛtelɔ, *ˈkɔmetɛlɔ	

- Likewise for pretonic vowels; post-tonic harmony is a prerequisite for pretonic harmony:

momentos	moˈmɛntɔ ∼ mɔˈmɛntɔ	'instants'
relojes	reˈlɔhɛ ∼ rɛˈlɔhɛ	'watches'
monederos	moneˈðɛɾɔ ∼ mɔnɛˈðɛɾɔ	'purses'
	*mɔneˈðɛɾɔ, *monɛˈðɛɾɔ	
recógelos	reˈkɔhelɔ ∼ reˈkɔhɛlɔ ∼ rɛˈkɔhɛlɔ	'pick them'
	*rɛˈkɔhelɔ	

Variable Harmony in Eastern Andalusian

- But high vowels do not undergo harmony:

crisis	ˈkɾisɪ	'crisis'
muchos	ˈmuʃɔ	'many'
ídolos	ˈiðolɔ ∼ ˈiðɔlɔ	'idols'
cojines	koˈhinɛ ∼ kɔˈhinɛ	'pillows'
cotillones	kotiˈʒɔnɛ ∼ kɔtiˈʒɔnɛ	'cotillions'

Core Constraints

- LICENSE([–ATR], $\acute{\sigma}$): assign $+1$ for each [–ATR] that coincides with $\acute{\sigma}$ and $+1$ for each additional syllable that [–ATR] appears in (Kaplan 2018b; Walker 2011).
  - Positive constraints require serialism (Kimper 2011).
- CRISPEDGE([–ATR], $\acute{\sigma}$, L): assign -1 for each syllable to the left of the stressed syllable with which it shares a [–ATR] feature (e.g. Ito & Mester 1999; Kaplan 2018c).
- *[–ATR]: assign -1 for each vowel bearing [–ATR].
- *[+hi, –ATR]: assign -1 for [ɪ, ʊ].

Add noise to the computation of harmony scores at various levels (Hayes 2017):

/rekóhelos/	LICENSE 11	CRISPEDGE 0.25	*[−ATR] 11	*H*
a. reˈkohelɔ			−1	−11
(☞) b. reˈkɔhelɔ	+2		−2	0
(☞) c. reˈkɔhɛlɔ	+3		−3	0
(☞) d. rɛˈkɔhɛlɔ	+4	−1	−4	−0.25
e. rɛˈkɔhelɔ	+3	−1	−3	−0.25

NHG: the Mechanics

Add noise to the computation of harmony scores at various levels (Hayes 2017): constraint ("classical NHG"),

/rekóhelos/	LICENSE 11 +.5	CRISPEDGE 0.25 −.2	*[−ATR] 11 +.3	*H*
a. reˈkohelɔ			−1	~~−11~~ → −11.3
(☞) b. reˈkɔhelɔ	+2		−2	~~0~~ → 0.4
(☞) c. reˈkɔhɛlɔ	+3		−3	~~0~~ → 0.6
(☞) d. rɛˈkɔhɛlɔ	+4	−1	−4	~~−0.25~~ → 0.75
e. rɛˈkɔhelɔ	+3	−1	−3	~~−0.25~~ → 0.55

Add noise to the computation of harmony scores at various levels (Hayes 2017): constraint ("classical NHG"), cell,

/rekóhelos/	LICENSE 11	CRISPEDGE 0.25	*[−ATR] 11	*H*	
a. reˈkohelɔ	+.9	−.8	−1 −.5	~~−11~~	→ −10.5
(☞) b. reˈkɔhelɔ	+2 +.5	−.7	−2 −.5	~~0~~	→ 2
(☞) c. reˈkɔhɛlɔ	+3 +0	+.6	−3 −.9	~~0~~	→ 2.7
(☞) d. rɛˈkɔhɛlɔ	+4 −.9	−1 +.7	−4 +.2	~~−0.25~~	→ −5.35
e. rɛˈkɔhelɔ	+3 +.2	−1 −.2	−3 +.6	~~−0.25~~	→ −1.25

NHG: the Mechanics

Add noise to the computation of harmony scores at various levels (Hayes 2017): constraint ("classical NHG"), cell, or candidate.

/rekóhelos/	LICENSE 11	CRISPEDGE 0.25	*[−ATR] 11	*H*	
a. re'kohelɔ			−1	~~−11~~ +.4	→ −10.6
(☞) b. re'kɔhelɔ	+2		−2	~~0~~ +.2	→ 0.2
(☞) c. re'kɔhɛlɔ	+3		−3	~~0~~ −.8	→ −0.8
(☞) d. rɛ'kɔhɛlɔ	+4	−1	−4	~~−0.25~~ −.3	→ −.55
e. rɛ'kɔhelɔ	+3	−1	−3	~~−0.25~~ .4	→ 0.65

Also MaxEnt (Goldwater & Johnson 2003)

Only classical NHG noise accounts for Eastern Andalusian in parallel NHG (Kaplan 2018a; Kaplan 2019); also in serial NHG...

Serial Versions of Constraint-Level Noise

1. Constant noise: weights are perturbed once at the outset, fixing their values for the whole derivation.

   Step 1: $w(C) + i$
   Step 2: $w(C) + i$

2. Variable Noise: weights are perturbed anew at each step in the derivation.

   Step 1: $w(C) + i$
   Step 2: $w(C) + j$

3. Cumulative variable noise: like variable noise, but the starting point for each step is the perturbed weights from the previous step.

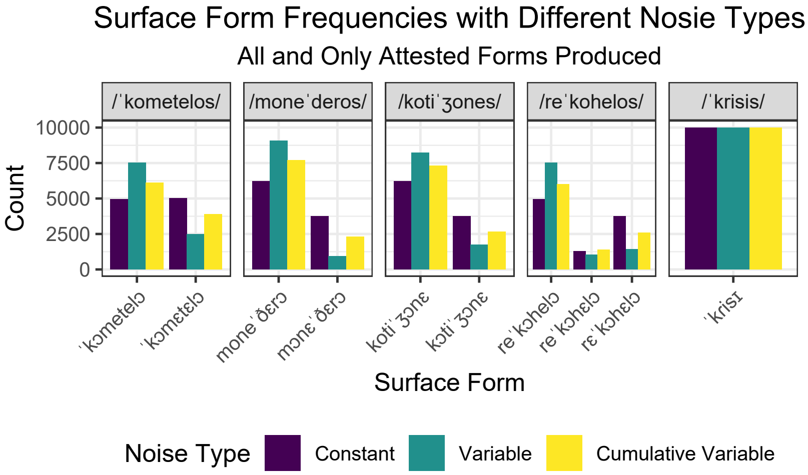
   Step 1: $w(C) + i$
   Step 2: $w(C) + i + j$

Simulations

- Existing software (OTsoft (Hayes, Tesar & Zuraw 2013), OT-Help (Staubs et al. 2010), e.g.) doesn't support serial NHG. (But OTSoft can help find constraint weights, as we'll see.)
- My own implementations, built in R (R Core Team 2020). Some details:
  - Noise was drawn from a normal distribution with mean of 0 and standard deviation of 1.
  - Negative weights were reverted to 0 (following Hayes (2017)).
  - In the event of tied winners, one is chosen at random.
  - Results from each implementation were aggregated over 10,000 iterations.
  - Weights supplied at the outset.

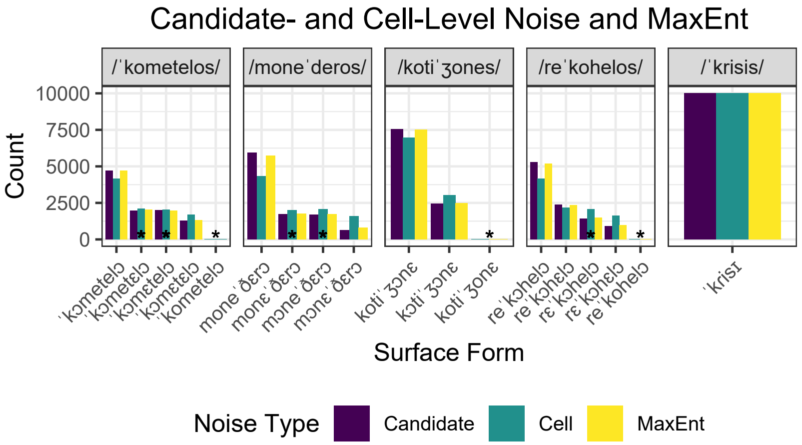
Simulations

- Fell-swoop harmony first.
- All three constraint-level noise simulations produce the licit outputs.
    - Different frequency predictions, but no way to assess them.
- Cell- and candidate-level noise and MaxEnt all overgenerate, unavoidably producing illicit candidates.



- Cumulative Variable: weights are doubled here to prevent accumulation of weights from subverting necessary dominance relationships.

Simulations



* = unattested; those with near-zero frequencies (8 tokens) all come from cell-level/MaxEnt.

A Gradual View of Harmony

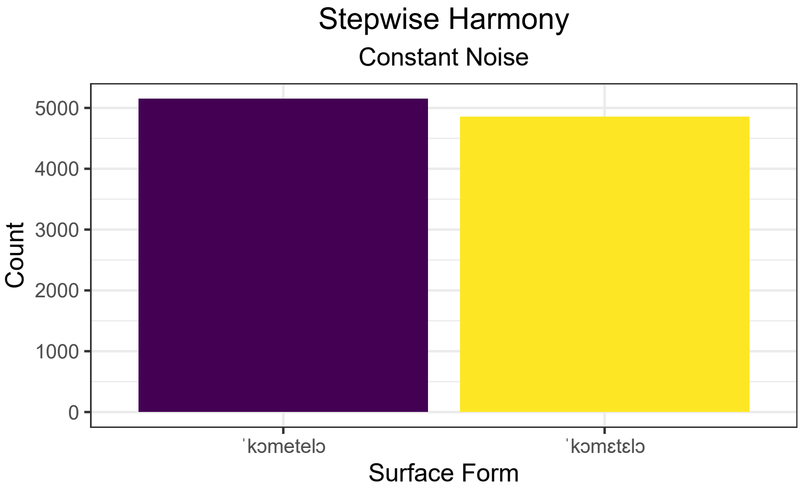
$$\text{re'kohelɔ} \xrightarrow{1} \text{re'kɔhelɔ} \xrightarrow{2} \text{re'kɔhɛlɔ} \xrightarrow{3} \text{rɛ'kɔhɛlɔ}$$

- How far down this path will we go?
- What if we treat feature changes, deletion, etc., as multi-step processes (e.g. McCarthy 2008)?

$$\text{re'kohelɔ} \xrightarrow{1} \text{re'kXhelɔ} \xrightarrow{2} \text{re'kɔhelɔ} \xrightarrow{3} \text{re'kɔhXlɔ} \xrightarrow{4} \text{re'kɔhɛlɔ} \xrightarrow{5} \text{rX'kɔhɛlɔ} \xrightarrow{6} \text{rɛ'kɔhɛlɔ}$$

- X = a V that's both [+ATR] and [–ATR]
- Now every other step is required. Can we ensure that Step 4 always happens if we choose to do Step 3, e.g.?
  - Yes: weight *DOUBLEASSOCIATION high enough that it will always trigger deletion of [+ATR], even after weights are perturbed.

A Stepwise View of Harmony



Summary: Eastern Andalusian

- Only constraint-level noise (all versions) provides a satisfactory analysis.
- The analysis succeeds with both fell-swoop and gradual harmony.
- Turning to Persian, focusing on constraint-level noise:
  - Test gradualism more fully
  - Match frequency data

Variable Hiatus Resolution in Persian

- Hiatus at morpheme boundaries is optionally resolved via epenthesis or deletion of the suffix-initial V (Ariyaee & Jurgec 2020):

/baba-emun/	→	[babaemun ∼ babamun ∼ babaʔemun]	'our dad'
/baba-eʃ	→	[babaeʃ ∼ babaʃ ∼ babaʔeʃ]	'his/her dad'
/baba-æm	→	[babaæm ∼ babam ∼ babaʔæm]	'my dad'

- If the suffix consists solely of the deleteable vowel, deletion is strongly disfavored (REALIZEMORPHEME (Kurisu 2001)):

/babae/ → [babae ∼ ???/*baba ∼ babaʔe] 'the dad'

- In the absence of hiatus Root + Suffix emerges unchanged:

/dæftær-emun/ → [dæftæremun] 'our office'

The Challenges

- Gradual deletion: can we ensure /babaemun/ → babaVmun → [babamun], not halting at *[babaVmun]?
- NoHiatus
  - Satisfied only by the final step in deletion, so it can't motivate the first step.
  - Let's assume NoHiatus penalizes consecutive *fully specified* vowels. Now [babaVmun] satisfies it.
- RealizeMorpheme
  - A standard view: as long as a morpheme has some phonological exponent, RealizeMorph is satisfied.
  - This will cause problems, and we'll revisit it later.

Constraints

- NoHiatus motivates epenthesis/deletion. (Penalizes [ae], not [aV].)
- RealizeMorpheme discourages deletion in /baba-e/.
- HavePlaceV penalizes placeless vowels, motivating V $\rightarrow \emptyset$.
- HavePlaceC penalizes [ʔ], hence penalizes epenthesis.
- Faithfulness:
  - MaxV (penalizes V $\rightarrow \emptyset$) & DepV (penalizes $\emptyset \rightarrow$ V)
  - MaxC & DepC
  - MaxVPlace (penalizes /e/ $\rightarrow$ V)
  - DepVPlace (penalizes V $\rightarrow$ [e])
- No MaxCPlace or DepCPlace: for simplicity, /ʔ/ $\rightarrow$ [t], e.g., not considered. Assume DepCPlace dominates everything.

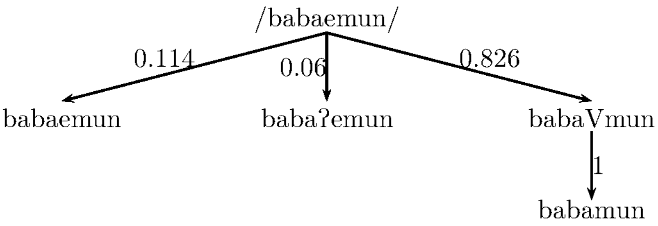
Variant Frequencies and Derivational Paths

Output variants inferred from graphs in Ariyaee & Jurgec (2020) using WebPlotDigitizer (Rohatgi 2020):

Surface Form	Target
/baba-e/	
babae	0.589
baba	0.080
babaʔe	0.331
/baba-emun/	
babaemun	0.114
babamun	0.826
babaʔemun	0.060
/dæftær-emun/	
dæftæremun	1.000

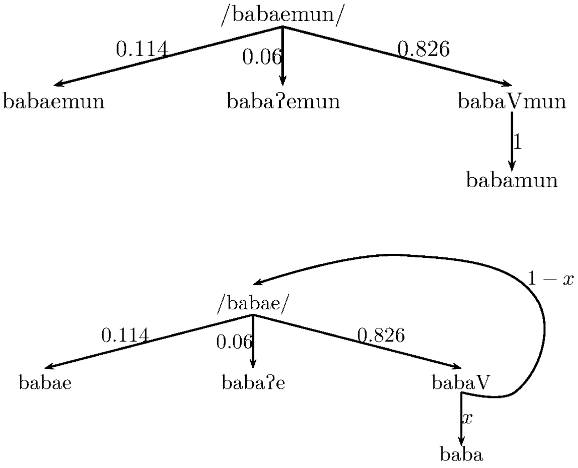
Variant Frequencies and Derivational Paths

- How do we get to these outputs? The simplest approach for /baba-emun/:



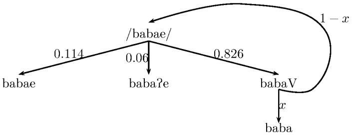
- /baba-e/: same paths, but different proportions due to REALIZEMORPH—the only constraint that distinguishes /baba-emun/'s derivations from /baba-e/'s.
- But REALIZEMORPH doesn't care about Step 1: it penalizes only /babaV/ $\rightarrow$ [baba], so the Step 1 proportions must match those for /baba-emun/.

Variant Frequencies and Derivational Paths



- What must the probability of /babaV/ → [baba] be to reduce the probability of deletion to .08?

Variant Frequencies and Derivational Paths



- Probability of convergence on [baba]:

$$0.826x + 0.826x(0.826(1-x)) + 0.826x(0.826(1-x))^2 + 0.826x(0.826(1-x))^3 + \ldots$$

- The infinite series $a + ar + ar^2 + ar^3 + \ldots = \frac{a}{1-r}$. Therefore:

$$\frac{0.826x}{1 - 0.826(1-x)} = .08$$

So $x \approx 0.018$

Deriving Weights

1. Create OTSoft file with each step in our derivations:

Input	Legal Outputs
/baba-e/	babae ∼ babaʔe ∼ babaV
babaʔe	babaʔe
babaV	baba ∼ babae
/baba-emun/	babaemun ∼ babaʔemun ∼ babaVmun
babaʔemun	babaʔemun
babaVmun	babamun
/dæftær-emun/	dæftæremun

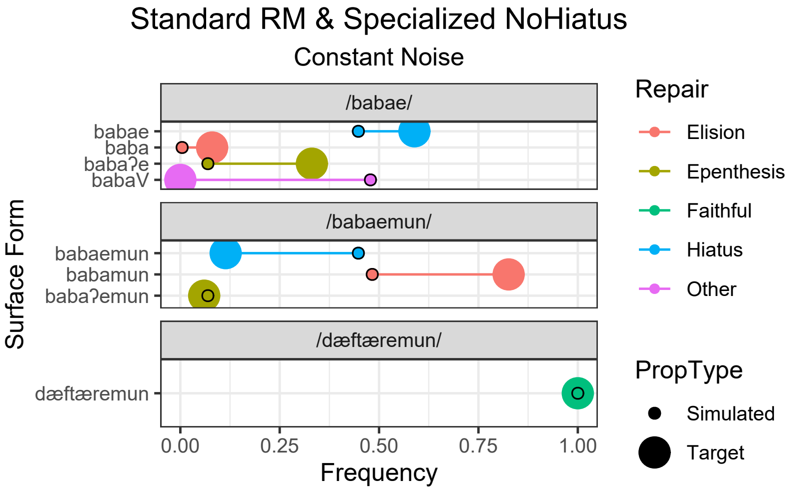
2. Probabilities for each candidate match what we saw before for /baba-e/ and /baba-emun/. For inputs with one possible output, that output's probability = 1. All other mappings = 0.

Deriving Weights

3. Submit file to OTSoft's NHG tool (premultiplicative constraint-level noise). The output:

Constraint	Weight	Remarks
NoHiatus	8.75	violated by [babaemun], not [babaVmun]
MaxVPlace	0.002	violated by /babaemun/ → [babaVmun]
DepC	6.78	violated by /babaemun/ → [babaʔemun]
MaxC	4	violated by /babaʔemun/ → [babaemun]
DepV	10	violated by /babamun/ → [babaVmun]
MaxV	1.15	violated by /babaVmun/ → [babamun]
HavePlaceV	8.31	violated by [babaVmun]
DepVPlace	0	violated by /babaVmun/ → [babaemun]
HavePlaceC	3.78	violated by [babaʔemun]
RealizeMorph	10.2	violated by [baba] but not [babV]

Results



Standard RM & Specialized NoHiatus
Constant Noise
/babae/
babae
baba
babaʔe
babaV
/babaemun/
babaemun
babamun
babaʔemun
/dæftæremun/
dæftæremun
Surface Form
0.00
0.25
0.50
0.75
1.00
Frequency
Repair
Elision
Epenthesis
Faithful
Hiatus
Other
PropType
Simulated
Target

Results

Target & Simulated Frequencies
Standard RM, Constant Noise

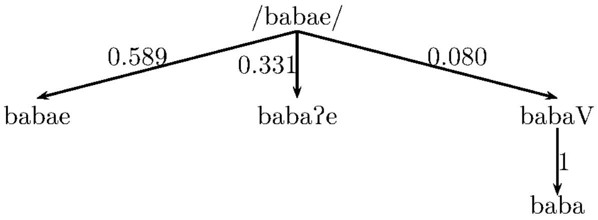
Surface Form	Target	Simulation
/baba-e/		
babae	0.589	0.4479
baba	0.080	0.0045
babaʔe	0.331	0.0693
babaV	0.000	0.4783
/baba-emun/		
babaemun	0.114	0.4479
babamun	0.826	0.4828
babaʔemun	0.060	0.0693
/dæftær-emun/		
dæftæremun	1.000	1.0000

Results

- Attempts to adjust weights to exclude [babaV] (e.g. by elevating HavePlaceV) also increased the likelihood of [baba], which quickly becomes the most common output for /babae/. (It should be the least common.)
- Perhaps there are weights that work, but I can't find them.

An Alterative RealizeMorph

- Alternative: match output frequencies at Step 1 as with /babaemun/:



- A new conception of RealizeMorph: in [babaV], the exponent of the suffix is a vowel lacking features. Maybe RM requires a *pronounceable* exponent for each morpheme.

An Alterative REALIZEMORPH

REALIZEMORPHEME: assign -1 for each morpheme that does not have a fully specified phonological exponent.

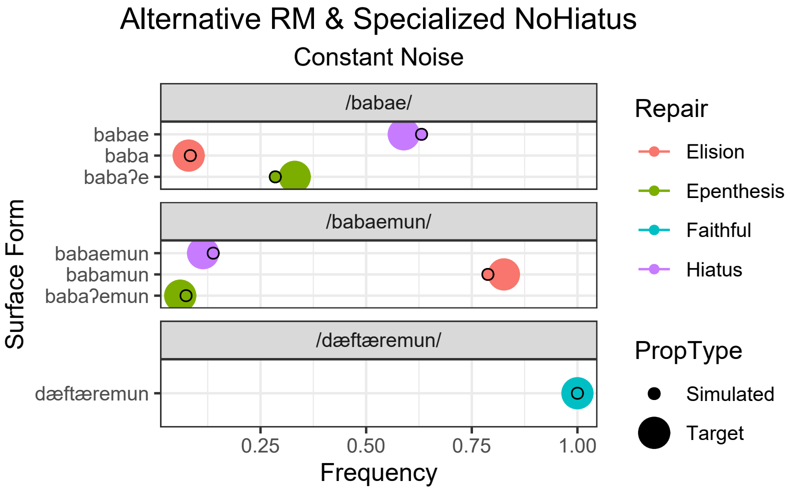
- What does "fully specified" mean? Not sure. . . let's assume that [V] doesn't cut it and worry about the details later.

An Alterative RealizeMorph

- Once again, using OTSoft to derive weights:

Constraint	New Weight	Old Weight
NoHiatus	11.2	8.75
MaxVPlace	2.74	0.002
DepC	7.03	6.78
MaxC	3	4
MaxV	0	1.15
HavePlaceV	6.74	8.31
DepVPlace	5	0
HavePlaceC	5.03	3.78
RealizeMorph	4.08	10.2

Results



Alternative RM & Specialized NoHiatus
Constant Noise
/babae/
babae
baba
babaʔe
/babaemun/
babaemun
babamun
babaʔemun
/dæftæremun/
dæftæremun
Surface Form
0.25
0.50
0.75
1.00
Frequency
Repair
Elision
Epenthesis
Faithful
Hiatus
PropType
Simulated
Target

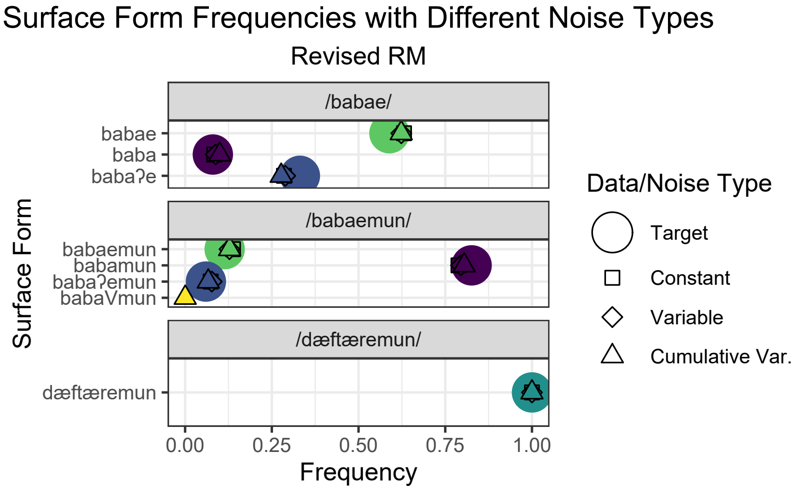
An Alterative RealizeMorph

Target & Simulated Frequencies
Alternative RM, Constant Noise

Surface Form	Target	Simulation	MaxEnt (A&J)
/baba-e/			
babae	0.589	0.6311	0.55
baba	0.080	0.0839	0.14
babaʔe	0.331	0.2850	0.31
/baba-emun/			
babaemun	0.114	0.1381	0.25
babamun	0.826	0.7881	0.61
babaʔemun	0.060	0.0738	0.14
/dæftær-emun/			
dæftæremun	1.000	1.0000	NA

Other Variants of Constraint-Level Noise

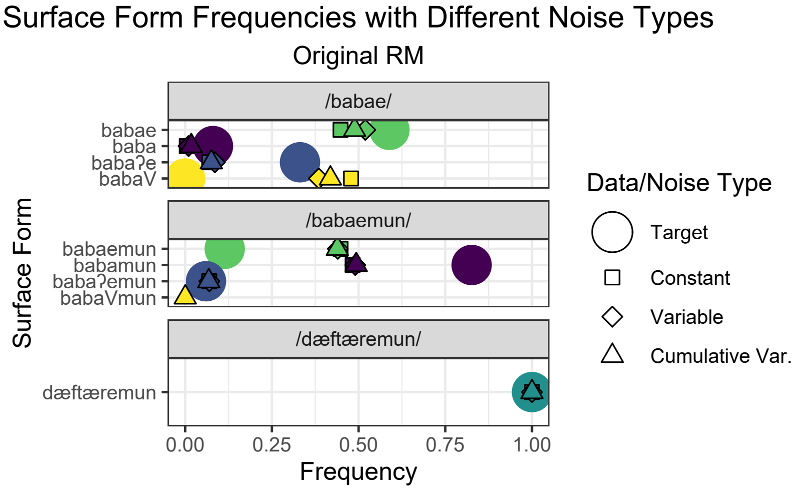
- The three versions of constraint-level noise are roughly similar:



- Cumulative Variable: doubling weights eliminates illicit outputs, but frequencies are less accurate.

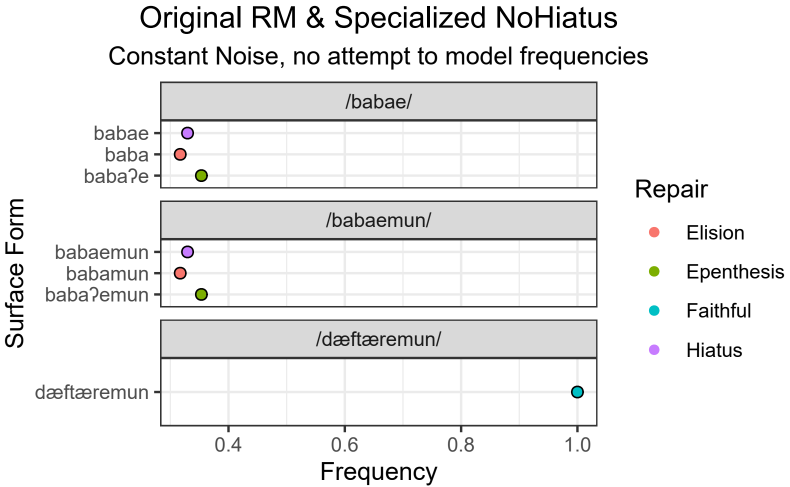
Other Variants of Constraint-Level Noise

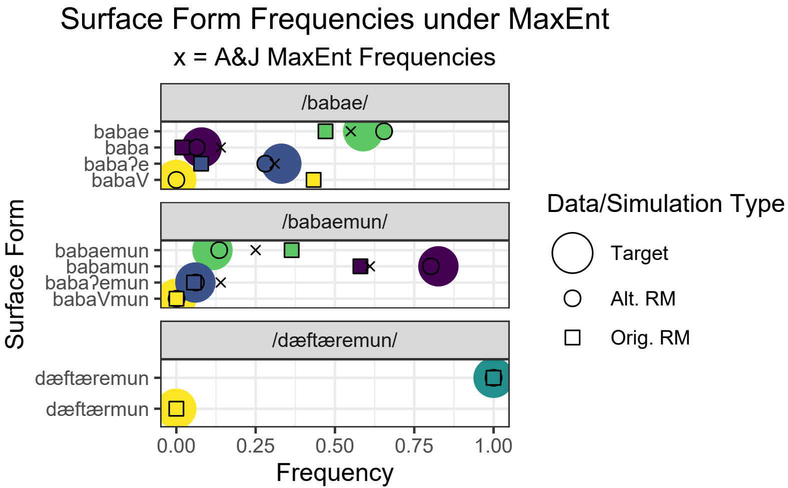
- And they all do poorly with the original REALIZEMORPH:



The Original REALIZEMORPH isn't all Bad

- Original RM can model outputs but not frequencies.
- Abandoning the frequency-matching effort (weights again derived via OTSoft):




Surface Form Frequencies under MaxEnt
x = A&J MaxEnt Frequencies
/babae/
babae
baba
babaʔe
babaV
/babaemun/
babaemun
babamun
babaʔemun
babaVmun
/dæftæremun/
dæftæremun
dæftærmun
Surface Form
0.00
0.25
0.50
0.75
1.00
Frequency
Data/Simulation Type
Target
Alt. RM
Orig. RM

Implications

- Modeling outputs is easy; matching frequencies is harder.
    - Frequencies require refinements to constraints.
- No frequencies are available for Eastern Andalusian—would the challenges presented by Persian reemerge in Eastern Andalusian if we had frequencies?
- Maybe, but there's an important difference between the two phenomena:
    - Persian: which path will we take? (Once that choice is made, the remainder of the derivation is deterministic.)
    - EA: how far down the path will we go? (Just the stressed syllable? Also post-tonic harmony? Also pretonic harmony?)

Implications

- Serial NHG is very similar to parallel NHG. 2 indications:
  - The version of parallel NHG that works for Eastern Andalusian is also the only version of serial NHG that works.
  - The weights that OTSoft provides when it thinks serial derivations are unrelated parallel evaluations hold up serially.

Remaining Issues

- Cell- and candidate-level noise?
- The simulations shown here use only premultiplicative noise: add noise to weight, then multiply by violations. What about post-multiplicative noise (for both Eastern Andalusian and Persian)?
- A better way to arrive at weights for Persian?

References I

Ariyaee, Koorosh & Peter Jurgec. 2020. Variable hiatus in Persian is affected by suffix length. Paper presented at AMP 2020, Santa Cruz, CA, September 19.

Flemming, Edward. 2017. Stochastic harmonic grammars as random utility models. Poster presented at AMP 2017.

Hayes, Bruce. 2017. Varieties of noisy HG. In Karen Jesney, Charlie O'Hara, Caitlin Smith & Rachel Walker (eds.), *Proceedings of AMP 2016*. 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/.

References II

Ito, Junko & Armin Mester. 1999. Realignment. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody-morphology interface*, 188–217. Cambridge, U.K.: Cambridge University Press.

Jesney, Karen. 2007. The locus of variation in weighted constraint grammars. Poster presented at the Workshop on Variation, Gradience and Frequency in Phonology. Stanford, CA: Stanford University. July 2007.

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.

References III

Kaplan, Aaron. 2018a. Noisy hg models of eastern andalusian harmony. Talk presented at Analyzing Typological Structure: From Categorical to Probabilistic Phonology, Stanford University, September 22.

Kaplan, Aaron. 2018b. Positional licensing, asymmetric trade-offs, and gradient constraints in Harmonic Grammar. *Phonology* 35(2). 247–286.

Kaplan, Aaron. 2018c. Asymmetric crisp edge. In Ryan Bennett, Adrian Brasoveanu, Dhyana Buckley, Nick Kalivoda, Shigeto Kawahara, Grant McGuire & Jaye Padgett (eds.), *Hana-bana: A festschrift for Junko Itô and Armin Mester*. Santa Cruz, CA: Department of Linguistics, University of California, Santa Cruz. https://itomestercelebration.sites.ucsc.edu/.

References IV

Kaplan, Aaron. 2019. Noisy HG models of Eastern Andalusian harmony. Talk presented at the *Sixteenth Old World Conference in Phonology*, University of Verona, Verona, Italy, January 18.

Kimper, Wendell. 2011. *Competing triggers: Transparency and opacity in vowel harmony*. Amherst, MA: University of Massachusets, Amherst PhD thesis.

Kurisu, Kazutaka. 2001. *The phonology of morpheme realization*. University of California, Santa Cruz PhD thesis.

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.

McCarthy, John J. 2008. The gradual path to cluster simplification. *Phonology* 25. 271–319.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rohatgi, Ankit. 2020. Webplotdigitizer: Version 4.4. https://automeris.io/WebPlotDigitizer.

Staubs, Robert, Michael Becker, Christopher Potts, Patrick Pratt, John J. McCarthy & Joe Pater. 2010. OT-help 2.0. Software package. Amherst, MA: University of Massachusets Amherst.

Walker, Rachel. 2011. *Vowel patterns in language*. New York: Cambridge University Press.

Zuraw, Kie & Bruce Hayes. 2017. Intersecting constraint families: An argument for Harmonic Grammar. *Language* 93(3). 497–548.