Class 11
Gradience in Phonology:
Gradient Symbolic Representations
05.02.2021

1 Gradience in Phonology

- Traditionally, the difference between phonology and phonetics has been defined, in part, by the opposition between **categoricality** and **gradience** (see, e.g., Ernestus 2011; see also the references cited by Hayes & Wilson 2008:382).
  → Phonology is *categorical*: it operates over a discrete set of (abstract) symbols.
  → Phonetics is *gradient*: it operates over a set of continuously valued (surface) features.
- Under this view, the mapping from phonology to phonetics — which are distinct (grammatical) modules — consists of the translation of abstract symbols into physical manifestations, via motor planning or the like (e.g. Zsiga 1997).

⋆ This approach relegates anything that looks gradient to the realm of phonetics, and demands that anything that is truly phonological is categorical.
  → In recent years, however, it has become clear that at least certain aspects of phonological knowledge are gradient in nature.

1.1 Frequency Matching

- One area of gradient phonological knowledge is of "**Frequency Matching**" (Hayes et al. 2009:826):

(1) "**LAW OF FREQUENCY MATCHING:** Speakers of languages with variable lexical patterns respond stochastically when tested on such patterns. Their responses aggregately match the lexical frequencies."

- Hayes et al. (2009:826) propose this Law based on the findings of Ernestus & Baayen (2003) and many others that speakers recreate gradient phonological patterns when *wug*-tested in the right way.

- For example, Ernestus & Baayen (2003) examine the way Dutch speakers generalize final voicing alternations.
- Dutch has a categorical process of final obstruent devoicing, which neutralizes a lexical contrast between voiced and voiceless obstruents.
- For each pair of voiced and voiceless obstruent, the relative lexical frequency of is distinct and (somewhat) abritrary:

(2) a. **Labial stops:** more non-alternators (/p/ → [p/_#], [p/_V]) than alternators (/b/ → [p/_#], [b/_V])
    b. **Labial fricatives:** more alternators (/v/ → [f/_#], [v/_V]) than non-alternators (/f/ → [f/_#], [f/_V])

- Ernestus & Baayen (2003) conducted a *wug*-test production experiment where Dutch speakers were given a form with a final voiceless obstruent and asked to produce a related form where that obstruent would be pre-vocalic.

→ Rather than categorically picking the most frequent option, or the least marked option, etc., the Dutch speakers produced both types of forms, in proportion to their distributions in the lexicon: (Figure from Hayes et al. 2009:826.)

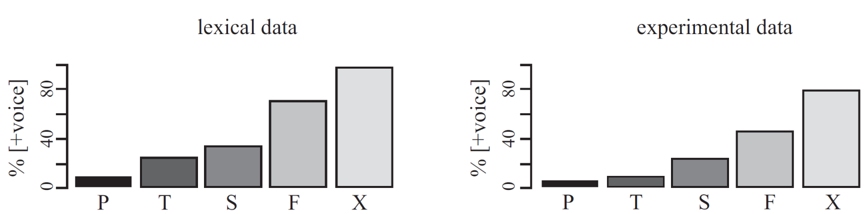

FIGURE 1. Frequency matching in Dutch voicing alternations: Ernestus & Baayen 2003.

- This kind of result has been replicated numerous times. It very clearly indicates that speakers have detailed statistical knowledge about variable patterns in their language.
  - ⋆ (It's not completely clear to me that this counts as grammatical knowledge...)

1.2 Gradient Phonotactics

- One other main area in which phonological knowledge is clearly gradient is in phonotactics.

- Hayes & Wilson (2008) find that a **Harmonic Grammar**-based learning model (MAXIMUM ENTROPY HARMONIC GRAMMAR) can learn constraints and associated weights that generate attested distributions of phonotactic sequences.
- This grammar in turn fits well to the phonotactic knowledge evidenced from nonce-word ratings tests (Albright 2009), indicating that speakers internalize distributional phonotactic generalizations from their lexicon (via constraints).

→ This is all gradient, because it distinguishes not just between grammatical and ungrammatical sequences, but sequences with continuously varying degrees of "good-ness" in between.

1.3 Gradience beyond variability?

- All of these kinds of "gradient" phonology have to do with morphologically/lexically variable patterns.
- To the extent that they are directly encoded in the grammar, this is primarily achieved through the use of **constraint weighting**, i.e. in **Harmonic Grammar** (Legendre, Miyata, & Smolensky 1990, Potts et al. 2010, *a.o*).

→ Yet the constraints themselves are still categorical, in that they refer to discrete symbolic representations.

- Researchers have also been exploring whether employing gradient **representations** can help explain any other sorts of phonological phenomena.
- One thrust of this research is to use gradient *phonetic* representations as the basis for (some kinds of) phonological constraints (e.g. Flemming 2001, Lionnet 2017, *a.o.*).

→ Another approach is to continue to use symbolic phonological representations, but introduce gradience into those representations.

2 Gradient Symbolic Representations

- Zimmermann (2018a,b, 2019) et seq. develops the theory of **Gradient Symbolic Representations (in the Output) [GSR(O)]** as a general theory of phonological representations.
  - ◦ This is based on Smolensky & Goldrick's (2016) and Rosen's (2016) framework of Gradient Symbolic Computation, an approach related to neural network models in cognitive science, implemented in Harmonic Grammar.

→ This approach is designed to deal with various sorts of exceptionality and other tricky problems that resist analysis by traditional means.

2.1 GSRO representations and constraints

- The main claim of GSRO is that all segments — in both the input and in the output — are indexed with a numerical **activity** value, which ranges (as a continuous variable) from **0** to **1**:

(3)
 a. "Normal" segments have full activation (**1**).
 b. Activity level **0** is equivalent to the absence of that segment.
 c. Segments may be specified as having an activity level in between **0** and **1** (and possibly also $>$ **1**?).

- Constraint violations are scaled with respect to activity levels, primarily in two ways:

(4)
 a. Faithfulness constraints measure changes in activity.
 b. Markedness constraints assign their violations in proportion to level of activity.

2.1.1 Constraints on (changes in) activity

- MAX and DEP assign violations based on the difference in activity level between input and output.

(5) Faithfulness constraints in GSRO (Zimmermann 2019:5)

a. **MAX:** Assign violation X for any segmental activity X in the input that is not present in the output.

b. **DEP:** Assign violation X for any segmental activity X in the output that is not present in the input.

- MAX penalizes **deletion** of *activity*.
  - (6a) deletes a fully active segment $/k_{1.0}/$, and so incurs 1.0 MAX violations.
  - (6b) deletes a partially active segment $/t_{0.5}/$, and so incurs 0.5 MAX violations.
  - (6c) reduces the activity of a fully segment $/k_{1.0}/$ by 0.5 to $[k_{0.5}]$, and so incurs 0.5 MAX violations.
  - → These violations are multiplied by the weight of the MAX constraint ($\boldsymbol{\mathcal{W}_1}$) to derive the total penalty assessed by that constraint for that candidate. This penalty contributes to that candidates *harmony* ($\boldsymbol{\mathcal{H}}$).

(6) **MAX violations**

$\mathcal{H}$ = *harmony*, $\mathcal{W}$ = *weight*		MAX $\mathcal{W}_1$	DEP $\mathcal{W}_2$	$\mathcal{H}$
a. **Deletion of *fully* active segment**	$/p_{1.0}a_{1.0}\mathbf{k_{1.0}}/ \rightarrow [p_{1.0}a_{1.0}]$	1.0		$\mathcal{W}_1$
b. **Deletion of *partially* active segment**	$/p_{1.0}a_{1.0}\mathbf{t_{0.5}}/ \rightarrow [p_{1.0}a_{1.0}]$	0.5		$.5\times\mathcal{W}_1$
c. ***Reduction* of fully active segment**	$/p_{1.0}a_{1.0}\mathbf{k_{1.0}}/ \rightarrow [p_{1.0}a_{1.0}k_{0.5}]$	0.5		$.5\times\mathcal{W}_1$

- DEP penalizes **insertion** of *activity*.
  - (7a) inserts a fully active segment $[i_{1.0}]$ which was not present at all in the input, and so incurs 1.0 DEP violations.
  - (7b) increases the activity of an underlyingly partially active segment $/t_{0.5}/$ by 0.5 to reach full activation in the output $[t_{1.0}]$, and so incurs 0.5 DEP violations.
  - (7c) inserts a partially active segment $[i_{0.5}]$, and so incurs 0.5 DEP violations.

(7) **DEP violations**

$\mathcal{H}$ = *harmony*, $\mathcal{W}$ = *weight*		MAX $\mathcal{W}_1$	DEP $\mathcal{W}_2$	$\mathcal{H}$
a. ***Insertion* of fully active segment**	$/p_{1.0}a_{1.0}k_{1.0}/ \rightarrow [p_{1.0}a_{1.0}k_{1.0}\mathbf{i_{1.0}}]$		1.0	$\mathcal{W}_2$
b. ***Activation* of partially active segment**	$/p_{1.0}a_{1.0}\mathbf{t_{0.5}}/ \rightarrow [p_{1.0}a_{1.0}t_{1.0}]$		0.5	$.5\times\mathcal{W}_2$
c. **Insertion of *partially* active segment**	$/p_{1.0}a_{1.0}k_{1.0}/ \rightarrow [p_{1.0}a_{1.0}k_{1.0}\mathbf{i_{0.5}}]$		0.5	$.5\times\mathcal{W}_2$

- The default state of all output segments is to have full activity (**1**). This is ensured by the constraint FULL:

(8) **FULL:** Assign violation 1–X for every output element with activity X. (Zimmermann 2019:5)

- This penalizes partially active segments in the output, based on their output activity level alone (⇒ *markedness*).
  - (9a) (= (6c)) reduced an underlyingly fully active segment $/k_{1.0}/$ by 0.5 to a partially active one $[k_{0.5}]$. This incurs 0.5 MAX violations, and also 0.5 FULL violations (1.0−0.5=0.5).
  - (9b) (= (7c)) inserted a partially active segment $[i_{0.5}]$. This incurs 0.5 DEP violations and 0.5 FULL violations.
  - (9c) faithfully realizes an underlyingly partially active segment $/t_{0.5}/$ as such. This incurs no faithfulness violations, but still does incur 0.5 FULL violations.
  - → (9a) and (9b) incur violations of multiple constraints. Their total harmony scores are the sum of their weighted constraint violations.

(9) **FULL violations**

		MAX $\mathcal{W}_1$	DEP $\mathcal{W}_2$	FULL $\mathcal{W}_3$	$\mathcal{H}$
a. ***Reduction* of fully active S**	$/p_{1.0}a_{1.0}\mathbf{k_{1.0}}/ \rightarrow [p_{1.0}a_{1.0}k_{0.5}]$	0.5		0.5	$.5\times\mathcal{W}_1 + .5\times\mathcal{W}_3$
b. **Insertion of *partially* active S**	$/p_{1.0}a_{1.0}k_{1.0}/ \rightarrow [p_{1.0}a_{1.0}k_{1.0}\mathbf{i_{0.5}}]$		0.5	0.5	$.5\times\mathcal{W}_2 + .5\times\mathcal{W}_3$
c. **Realization of *partially* active S**	$/p_{1.0}a_{1.0}\mathbf{t_{0.5}}/ \rightarrow [p_{1.0}a_{1.0}t_{0.5}]$			0.5	$.5\times\mathcal{W}_3$

2.1.2 Scaled markedness constraints

- In Smolensky & Goldrick's (2016) original proposal, partially active segments were simply not allowed in outputs.
  → So the outputs in (9) would be ruled out by GEN.
- Zimmermann (2019), however, proposes that these are licit output structures, and that permitting them allows for desirable effects with respect to markedness constraints.

(10) **NOCODA violations**

		FULL $\mathcal{W}_3$	NOCODA $\mathcal{W}_4$	$\mathcal{H}$
a. ***Fully* active coda segment**	$/\mathrm{p}_{1.0}\mathrm{a}_{1.0}\mathbf{k}_{\mathbf{1.0}}/ \rightarrow [\mathrm{p}_{1.0}\mathrm{a}_{1.0}\mathrm{k}_{1.0}]$		1.0	$\mathcal{W}_4$
b. ***Partially* active coda segment**	$/\mathrm{p}_{1.0}\mathrm{a}_{1.0}\mathbf{t}_{\mathbf{0.5}}/ \rightarrow [\mathrm{p}_{1.0}\mathrm{a}_{1.0}\mathrm{t}_{0.5}]$	0.5	0.5	$.5\times\mathcal{W}_3 + .5\times\mathcal{W}_4$

- Partially active output segments thus incur lower markedness penalties than corresponding fully active segments.

⋆ Under the right constraint weighting, this can result in partially active segments being resistant to markedness repair.

2.2 "Ghost" segments

- Zimmermann (2019) calls underlyingly partially active segments "**ghost segments**", because they can appear and disappear in unexpected (i.e. exceptional) ways.
- Zimmermann (2019:2, 6) identifies two distinct kinds of ghosts, based as determined by their underlying activity level (assuming fully-active output segments and equal weighting of MAX and DEP):

(11) **Appearing ghosts:** Segments that appear just in case that resolves a markedness problem; their *default state is to be unrealized.* $\boxed{\text{UNDERLYING ACTIVITY } n\text{: } 0.0 < n < 0.5}$

→ In the absence of markedness problems, it will be *more* costly to raise it up to full activity (DEP violations > 0.5) then to lower it to zero activity (MAX violation < 0.5).

(12) **Disappearing ghosts:** Segments that *dis*appear just in case that resolves a markedness problem; their *default state is to be realized.* $\boxed{\text{UNDERLYING ACTIVITY } n\text{: } 0.5 < n < 1.0}$

→ In the absence of markedness problems, it will be *less* costly to raise it up to full activity (DEP violations < 0.5) then to lower it to zero activity (MAX violation > 0.5).

- Zimmermann (2019) explores the typology of ghost segments in GSRO. These are essentially segments that do or don't surface in ways that are exceptional from the perspective of the rest of the language's phonology.
- She explores two main case studies: Welsh (Celtic) and Ahousaht Nuu-chah-nulth (Salishan).

(13) a. **Welsh:** There are multiple different ghost segments with different activity levels, but when they surface, they have full activation.

b. **Nuu-chah-nulth:** All ghosts have the same activity level, and they maintain that activity level in the output; this explains differential behavior with respect to markedness.

3 Welsh ghosts

3.1 Data

- The preposition 'with' alternates between a consonant-final form *gudag*, which appears when it precedes a vowel (14a), and *guda*, which appears when it precedes a consonant (14b).

(14) Welsh *guda(g)* 'with' (Zimmermann 2019:3, from Hannahs & Tallerman 2006:798, 810)

a. **Pre-vocalic** → *gudag* (**guda*)			b. **Pre-consonantal** → *guda* (**gudag*)		
gudag eraill	[g.uda.**g**-e.raiɬ]	'with others'	*guda gwên*	[gu.da.-gweːn]	'with a smile'
gudag un	[gu.da.**g**-un]	'with one'	*guda tri*	[gu.da.-tri]	'with three'

→ This distribution ensures perfect consonant/vowel alternation, maximizing satisfaction of the syllabic markedness constraints NOHIATUS (15a) and NOCODA (15b).

(15) a. **NOHIATUS (*V.V):** Assign a violation for each sequence of heterosyllabic vowels (hiatus).
b. **NOCODA:** Assign a violation for each syllable with a coda consonant.

⋆ We can characterize *guda(g)*'s behavior as:

(16) a. The [g] surfaces whenever its absence would create hiatus (14a).
b. The [g] doesn't surface if it would create a coda (14b).

- We find similar, though more complex behavior with the definite article.
  - In phrase-initial position, it surfaces as [ər] before a vowel (17a) and as [ə] before a consonant (17b).
  - When it follows a vowel within its phrase, such as the preposition *o* 'from', it always surfaces as [r], whether the following segment is a vowel (17c) or a consonant (17d).

(17) Welsh definite article *yr* ∼ *y* ∼ *r* [ər ∼ ə ∼ r] (Zimmermann 2019:3, from Hannahs & Tallerman 2006:782–783)

	Phrase-initial			Post-vocalic		
Pre-vocalic	a. *yr afon*	[**ə.r**-a.von]	'the river'	c. *o'r afon*	[o-.**r**-a.von]	'from the river'
Pre-consonantal	b. *y llyfr*	[**ə**-.ɬəvr]	'the book'	d. *o'r llyfr*	[o-**r**-.ɬəvr]	'from the book'

⋆ It is clear that the distribution is responding to some of the same markedness pressures as *guda(g)*, but it's not exactly the same. (We'll come back to the details.)

- Despite the fact that both *guda(g)* and *yr* clearly respond to the following context, when they appear together, they always surface as *gyda'r* [gəda-r-].
  - When the following segment is a vowel (18a), this creates perfect C/V alternation.
  - When the following segment is a consonant (18b), this leads to a coda; but we know from (17d) that the definite article can sometimes do this.

(18) Combining the two: *gyda'r* (Zimmermann 2019:3, from Hannahs & Tallerman 2006:784–785)

a.	**Pre-vocalic**	*gyda'r offer*	[gəda.-**r**-o.fer]	'with the equipment'	(*[gəda.**g**-**ə.r**-o.fer])
b.	**Pre-consonantal**	*gyda'r nod*	[gəda.-**r**.-nod]	'with the aim'	(*[gəda.**g**-**ə**(**r**).-nod])

- The problem is, it seems like there's better options in both cases.

(19) a. In the pre-vocalic environment, you could continue to get C/V alternation if you realized *all* of the segments: *[gəda.**g**-**ə.r**-o.fer].
b. In the pre-consonantal environment, you could get an improved C/V alternation if you realized *the other two* segments: *[gəda.**g**-**ə**.-nod].
c. Alternatively, realizing all three segments would yield the same syllable structure (one coda) as the attested form: *[gəda.**g**-**ər**.-nod].

- This doesn't make any sense if all of the segments *want* to surface, which we would expect because of MAX.
  ⇒ Zimmermann's (2019) insight: only some of the segments **want to surface**, while others **don't want to surface**.

(20) a. The /g/ in *guda(g)* **doesn't want to surface**, but it will begrudgingly do so if it's needed to avoid hiatus.
↪ It's an "**appearing ghost**"
b. The definite article's segments **want** to appear, but they can delete if it solves markedness problems.
↪ They're "***dis*appearing ghosts**".

- These are ghost (≈ exceptional) segments because they do not conform to the normal behavior of segments in the language. Other segments don't get deleted or inserted to fix syllable structure problems.

3.2 Analysis

- Zimmermann's (2019) analysis of Welsh is as follows:

(21) **Ghost activities**

a. The final /g/ in *guda(g)* underlyingly has **0.2 activity**: $/g_{1.0}u_{1.0}d_{1.0}a_{1.0}\mathbf{g_{0.2}}/$

b. Both segments of the definite article underlyingly have **0.6 activity**: $/\mathbf{ə_{0.6}r_{0.6}}/$

(22) **Constraint weights:**

a. Weight of FULL = 100 (or any very high number)

b. Weight of MAX = 10

c. Weight of DEP = 10

d. Weight of NOHIATUS = 7

e. Weight of NOCODA = 5

⋆ Note that these are **not** the *unique* combinations of weights and activities that could derive the distribution.
  - Since both weights and activities are continuous variables, there should (I think) be an infinite number of solutions to the problem.

- In Zimmermann's analysis, all the output segments are assumed to have full activity, enforced by assigning a very high weight of FULL. I will generally ignore output candidates with partially-activated output segments.

3.2.1 No repairs with non-ghosts

- The fact that MAX and DEP both outweigh both of the markedness constraints means that fully-active segments will not be deleted or inserted to repair those structures:

(23) **No repair for fully-active codas**

$/a_{1.0}v_{1.0}o_{1.0}n_{1.0}/$	MAX 10	DEP 10	NOHIATUS 7	NOCODA 5	Harmony [$\mathcal{H}$]
a. ☞ $a_{1.0}v_{1.0}o_{1.0}n_{1.0}$				1	5
b. $a_{1.0}v_{1.0}o_{1.0}$	1.0				10
c. $a_{1.0}v_{1.0}o_{1.0}n_{1.0}ə_{1.0}$		1.0			10

(24) **No repair for fully-active hiatus**

$/o_{1.0}\ \ a_{1.0}v_{1.0}o_{1.0}n_{1.0}/$	MAX 10	DEP 10	NOHIATUS 7	NOCODA 5	$\mathcal{H}$
a. ☞ $o_{1.0}.a_{1.0}v_{1.0}o_{1.0}n_{1.0}$			1	1	12
b. $o_{1.0}v_{1.0}o_{1.0}n_{1.0}$	1.0			1	15
c. $o_{1.0}ʔ_{1.0}a_{1.0}v_{1.0}o_{1.0}n_{1.0}$		1.0		1	15

- Deleting a normal consonant ($/n_{1.0}/$) to get rid of a coda (23b), or a normal vowel ($/a_{1.0}/$) to get rid of hiatus (24b), will incur a greater penalty ($1.0 \times 10 = 10$) than allowing the corresponding marked structure to appear ($1 \times 5 = 5$ for codas; $1 \times 7 = 7$ for hiatus).
- The same goes for inserting a full vowel ($[ə_{1.0}]$ in (23c)) or consonant ($[ʔ_{1.0}]$ in (24c)).

- Inserting a partially active segment, or partially deactivating an underlying full segment, which would change the penalty scores, is not allowed because of high-weighted FULL.

3.2.2 *guda(g)*

- The "default" behavior for *guda(g)*'s final $/g_{0.2}/$ is evident in pre-consonantal position (25).
  - Since its underlying activity is < 0.5, it is more costly to fully activate the $/g_{0.2}/$ ($0.8 \times 10 = 8$ penalty for DEP) in candidate (25a) than to delete it ($0.2 \times 10 = 2$ penalty for MAX) in candidate (25b).
  - Additionally, surfacing in this position incurs a penalty for NOCODA.

(25) **Non-surfacing of /$\mathbf{g_{0.2}}$/ in pre-consonantal position**

/guda$\mathbf{g_{0.2}}$ tri/		MAX 10	DEP 10	NOCODA 5	$\mathcal{H}$
a.	gu.dag$_{1.0}$.-tri		0.8	1	13
b. ☞	gu.da.-tri	0.2			2

- We derive the "appearing ghost" behavior of *guda(g)* from a "gang effect".
  - In pre-vocalic position, the deletion candidate (26b) now incurs a penalty of 7 from NOHIATUS.
  - While it is still more costly to activate than delete, the combined penalty of deleting the segment and having a hiatus (26b) is *greater* than the penalty for just activating (26a).
  - → Thus, MAX and NOHIATUS *gang up* on DEP to prefer (26a), even though neither violation would be strong enough on its own to trigger activation.

(26) **Surfacing of /$\mathbf{g_{0.2}}$/ in pre-vocalic position**

/guda$_{1.0}$$\mathbf{g_{0.2}}$ un/		FULL 100	MAX 10	DEP 10	NOHIATUS 7	$\mathcal{H}$
a. ☞	gu.da.g$_{1.0}$-un			0.8		8
b.	gu.da.-un		0.2		1	9
c.	gu.d-un		1.2			12
d.	gu.da.g$_{0.2}$-un	0.8				80

- Additionally deleting a full vowel to avoid the hiatus violation (26c) is worse than activating the /g$_{0.2}$/, because it incurs an additional penalty of 10 from MAX.
- Realizing the /g$_{0.2}$/ faithfully as a partially-active segment (26d), which satisfies all other markedness and faithfulness constraints fully is in principle allowed.
  - → However, it is ruled out by very high-weighted FULL, which penalizes partially-active output segments.

3.2.3 *yr*

- The choice of candidate (27a), which realizes both segments of *yr*, over candidate (27b), which realizes only the /r/, shows that (at least) the /ə/ must be a "disappearing ghost".
  - i.e., its default state, which we can see when no markedness problems are at stake, is that it prefers to appear.
  - This is instantiated by its activity of 0.6, which falls in the range of $0.5 < n < 1.0$.
  - → It is thus more costly to delete it (MAX penalty of 6) than to activate it (DEP penalty of 4).

(27) **Full surfacing of /$\mathbf{ə_{0.6}r_{0.6}}$/ in pre-vocalic position**

/$\mathbf{ə_{0.6}r_{0.6}}$ avon/		MAX 10	DEP 10	NOHIATUS 7	NOCODA 5	$\mathcal{H}$
a. ☞	ə$_{1.0}$.r$_{1.0}$-a.vo...		0.8 (ə + r)			8
b.	r$_{1.0}$-a.vo...	0.6 (ə)	0.4 (r)			10
c.	ə$_{1.0}$.-a.vo...	0.6 (r)	0.4 (ə)	1		17
d.	a.vo...	1.2 (ə + r)				12

- The fact that the /r$_{0.6}$/ deletes in (28) shows that the weight of NOCODA is enough to flip the preference for activation (28a) to deletion (28b); i.e., even though its a bigger change to delete than activate, activating in this case incurs an additional markedness penalty from NOCODA.

(28) **Surfacing of /$\mathbf{ə_{0.6}}$/ in pre-consonantal position**

/$\mathbf{ə_{0.6}r_{0.6}}$ ɬəvr/		MAX 10	DEP 10	NOHIATUS 7	NOCODA 5	$\mathcal{H}$
a.	ə$_{1.0}$r$_{1.0}$.-ɬə...		0.8 (ə + r)		1	13
b. ☞	ə$_{1.0}$.-ɬə...	0.6 (ə)	0.4 (r)			10

- The deletion of the /ə$_{0.6}$/ in (29b) shows the exact same thing with respect to NoHiatus.

(29) **Surfacing of /r$_{0.6}$/ in post-vocalic, pre-vocalic position**

/**o ə$_{0.6}$r$_{0.6}$** avon/			Max 10	Dep 10	NoHiatus 7	NoCoda 5	$\mathcal{H}$
a.		o.-ə$_{1.0}$.r$_{1.0}$-a.vo...		0.8 (ə + r)	1		15
b.	☞	o.-r$_{1.0}$-a.vo...	0.6 (ə)	0.4 (r)			10
c.		o.-ə$_{1.0}$.-a.vo...	0.6 (r)	0.4 (ə)	2		24

- The tricky case is (30) where the /r$_{0.6}$/ surfaces in coda position between a preceding vowel and following consonant.
- Given what we've seen thus far, we would expect both segments to delete (30d):
  - The /ə$_{0.6}$/ should delete to avoid hiatus.
  - The /r$_{0.6}$/ should delete to avoid a coda.

(30) **Surfacing of /r$_{0.6}$/ in post-vocalic, pre-consonantal position** [to be refined]

/o **ə$_{0.6}$r$_{0.6}$** ɬəvr/			Max 10	Dep 10	NoHiatus 7	NoCoda 5	$\mathcal{H}$
a.		o.-ə$_{1.0}$r$_{1.0}$.-ɬə...		0.8 (ə + r)	1	1	20
b.	☹	o-r$_{1.0}$.-ɬə...	0.6 (ə)	0.4 (r)		1	15
c.		o.-ə$_{1.0}$.-ɬə...	0.6 (r)	0.4 (ə)	1		17
d.	💣	o.-ɬə...	1.2 (ə + r)				12

- Since this candidate uniquely deletes the entire content of the morpheme, we can rule it out by invoking a constraint against *completely deleting morphemes* (Zimmermann 2019:8).
- This constraint has been proposed as a solution for various other problems, and is typically called Realize Morpheme:

(31) **RealizeMorpheme [RM]:** Assign 1 violation for every morpheme in the input with which no segment in the output is associated. (Kurisu 2001; cf. van Oostendorp 2007)

- If RM is highly weighted, this will eliminate the problematic double-deletion candidate (32d). The evaluation will now select the best candidate that deletes only one of the segments.
  - This is determined by the relative weighting of the markedness constraints.
  - Evidently, ending up with a coda is better than ending up with hiatus.
  - This means that NoHiatus must be weighted higher than NoCoda. This selects (32b) over (32c).

(32) **Surfacing of /r$_{0.6}$/ in post-vocalic, pre-consonantal position**

/o **ə$_{0.6}$r$_{0.6}$** ɬəvr/			RM 100	Max 10	Dep 10	NoHiatus 7	NoCoda 5	$\mathcal{H}$
a.		o.-ə$_{1.0}$r$_{1.0}$.-ɬə...			0.8 (ə + r)	1	1	20
b.	☞	o-r$_{1.0}$.-ɬə...		0.6 (ə)	0.4 (r)		1	15
c.		o.-ə$_{1.0}$.-ɬə...		0.6 (r)	0.4 (ə)	1		17
d.		o.-ɬə...	1	1.2 (ə + r)				112

3.2.4 *guda(g) + yr*

[There's vowel reduction in the first syllable of the preposition here, which I'll ignore.]

- The interaction that derives consistent *guda'r* is complicated to tease apart, but it works.

- For the pre-vocalic case (33), there are two outputs that fully satisfy markedness:
  - (33a), which realizes all the segments
  - (33d), which realizes just /$\mathrm{r}_{0.6}$/
- They share the 0.4 DEP violation from the surfacing of /$\mathrm{r}_{0.6}$/, so that can be factored out.
- The determination is then made by the combined difference in activation of /$\mathrm{g}_{0.2}$/ and /$\mathrm{ə}_{0.6}$/ from **0** vs. from **1**.
  - /$\mathrm{g}_{0.2}$/ is closer to **0** (it's preferred state) than /$\mathrm{əg}_{0.6}$/ is to **1** (it's preferred state).
  - Therefore, deleting both (33d) (i.e. bringing their activation to **0**) will be cheaper [$(0.2 + 0.6) \times 10 = 8$] than activating both (33a) (i.e. bringing their activation to **1**) [$(0.8 + 0.4) \times 10 = 12$].

(33) ***guda(g) + yr* → *guda'r* in pre-vocalic position**

/guda$\mathbf{g_{0.2}}$ $\mathbf{ə_{0.6}r_{0.6}}$ avon/		MAX 10	DEP 10	NOHIATUS 7	NOCODA 5	$\mathcal{H}$
a.	gu.da.$\mathrm{g}_{1.0}$-$\mathrm{ə}_{1.0}$.$\mathrm{r}_{1.0}$-a.vo...		1.6 (g + ə + r)			16
b.	gu.da.-$\mathrm{ə}_{1.0}$.$\mathrm{r}_{1.0}$-a.vo...	0.2 (g)	0.8 (ə + r)	1		17
c.	gu.da.$\mathrm{g}_{1.0}$-$\mathrm{ə}_{1.0}$.-a.vo...	0.6 (r)	1.2 (g + ə)	1		25
d. ☞	gu.da.-$\mathrm{r}_{1.0}$-a.vo...	0.8 (g + ə)	0.4 (r)			12

- The same logic will prefer realization of just /$\mathrm{r}_{0.6}$/ (34d) to realization of all three (34a) in the pre-consonantal case, where both candidates additionally share a NOCODA violation.
- Here, the phonotactically optimal candidate is (34c), which realizes just /$\mathrm{g}_{0.2}$/ and /$\mathrm{ə}_{0.6}$/.
- While more is going on in this case, it boils down to the idea that raising /$\mathrm{g}_{0.2}$/'s activity level up to **1** is so costly that it ends up outweighing the cost of a NOCODA violation.

(34) ***guda(g) + yr* → *guda'r* in pre-consonantal position**

/guda$\mathbf{g_{0.2}}$ $\mathbf{ə_{0.6}r_{0.6}}$ ɬəvr/		MAX 10	DEP 10	NOHIATUS 7	NOCODA 5	$\mathcal{H}$
a.	gu.da.$\mathrm{g}_{1.0}$-$\mathrm{ə}_{1.0}\mathrm{r}_{1.0}$.-ɬə...		1.6 (g + ə + r)		1	21
b.	gu.da.-$\mathrm{ə}_{1.0}\mathrm{r}_{1.0}$.-ɬə...	0.2 (g)	0.8 (ə + r)	1	1	22
c.	gu.da.$\mathrm{g}_{1.0}$-$\mathrm{ə}_{1.0}$.-ɬə...	0.6 (r)	1.2 (g + ə)			18
d. ☞	gu.da.-$\mathrm{r}_{1.0}$-ɬə...	0.8 (g + ə)	0.4 (r)		1	17

3.3 Conclusions from Welsh

- Zimmermann (2019) shows that the complex interactions of *guda(g)* and *yr* in different phonological contexts can be accounted for in a weighted constraint model operating on Gradient Symbolic Representations.
  - Though Brinkerhoff (2019) has a recent analysis based on priority of allomorph selection Mascaró (2007) that seems to capture the data as well.
- Because of the differing behaviors of the two morphemes, she concludes that languages must be allowed to have different kinds of "ghosts".
  → This is difficult to generate with other approaches to exceptionality, like indexed-constraint theory (Pater 2009) or approaches relying on featural faithfulness.

4 Partially activated outputs in GSRO

4.1 Nuu-chah-nulth ghosts and gradient markedness

4.1.1 Data

- Nuu-chah-nulth has ghost consonants at the beginning of two types of morphemes:
  - /**-C👻V**/ (35a,c)
  - /**-C👻CV**/ (35b,d)
- Both have the same surface distribution: C👻 appears after vowels (35a,b) but not after consonants (35c,d).

(35) Nuu-chah-nulth ghosts (Zimmermann 2019:4, from Kim 2003:178)

			Post-vocalic		Post-consonantal
-C👻V	/-**q**👻umɬ/ 'round'	a.	[ʔat͡ɬa.-**q**👻umɬ] 'two dollars'	c.	[t͡ɬi.s-umɬ] 'sth. white and round'
-C👻CV	/(ʔu)-**k**👻ɬaː/ 'to be called'	b.	[ʔu-**k**👻.ɬaː-siʃ] 'my name is'	d.	[kʷis.-ɬaː-k'uk-ʔiʃ] 'it seems like he has a different name'

- Here's the problem:
  - When the **-C👻V** disappears in post-consonantal position (35b), this should be to satisfy NoCoda.
  - But when the **-C👻CV** appears in post-vocalic position (35c), it creates a coda, and thus a NoCoda violation.

→ Why doesn't the **-C👻CV** also disappear in post-vocalic position?

- Answer: a NoCoda violation from a partially-active consonant is not enough to trigger deletion; only a NoCoda violation from a fully active consonant.

4.1.2 Analysis

- **-C👻CV** is realized after a vowel because the ghost can be in the coda.
  - The combined weight of 0.5 Full violations plus (crucially) 0.5 NoCoda violations (36a) is *less* than the weight of 0.5 Max violations (36b), so (36a) is selected.

(36)

/ʔu-$\mathbf{k_{0.5}}$ɬaː/	Max 20	Full 12	NoCoda 7	$\mathcal{H}$
a. ☞ ʔu-$\mathbf{k_{0.5}}$.ɬaː		0.5	0.5	9.5
b. ʔu-ɬaː	0.5			10

- **-C👻V** not realized after a consonant because that would require a full consonant (the one belonging to the stem) to be in the coda.
  - In contrast with the previous case, now the combined weight of 0.5 Full violations plus a whole NoCoda violations (37a) is *greater* than the weight of 0.5 Max violations (37b), so (37b) is selected.

(37)

/t͡ɬis-$\mathbf{q_{0.5}}$umɬ/	Max 20	Full 12	NoCoda 7	$\mathcal{H}$
a. t͡ɬis.-$\mathbf{q_{0.5}}$u...		0.5	1	13
b. ☞ t͡ɬi.s-u...	0.5			10

- As long as Dep is much higher-weighted than Full, it will be preferable to have a partially-active segment surface than to boost its activity to **1**.

4.2 Gradient Backness in Uyghur Vowel Harmony

- Zimmermann (2019:9) states that, in Nuu-chah-nulth, there is no phonetic distinction between a partially active surface consonant and a fully active surface consonant.
  - She says there must be some phonetic/post-phonological mechanism that neutralizes the distinction.
- But she suggests that this may not be the case for all languages (i.e. some might maintain phonetic differences based on activity), and that Uyghur might be such a case:

- McCollum (2019) shows that Uyghur backness harmony must simultaneously be:
  - Phonological — because harmonized vowels can trigger allophonic processes, and
  - Gradient — because phonetic backness decreases gradiently across the harmonizing span

→ This could be interpreted as another case of partially active output segments:
  - This kind of gradience could be the result of partial activation of the feature [+back] on harmonizing vowels, if that activation could be made to decrease proportionally based on distance from the trigger of harmony.
- Notably, this partial activity would be *derived*, not simply retained from the input.

References

Albright, Adam. 2009. Feature-Based Generalisation as a Source of Gradient Acceptability. *Phonology* 26(1):9–41.

Brinkerhoff, Mykel Loren. 2019. Optimality Theory with Lexical Selection Needs Multiple Priorities: A Case Study in Welsh Allomorphy. *Proceedings of the Linguistic Society of America* 4(24):1–14.

Ernestus, Miriam & R. Harald Baayen. 2003. Predicting the Unpredictable: Interpreting Neutralized Segments in Dutch. *Language* 79(1):5–38.

Ernestus, Mirjam. 2011. Gradience and Categoricality in Phonological Theory. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, 1–22. John Wiley & Sons, Ltd.

Flemming, Edward. 2001. Scalar and Categorical Phenomena in a Unified Model of Phonetics and Phonology. *Phonology* 18(1):7–44.

Hannahs, S. J. & Maggie Tallerman. 2006. At the Interface: Selection of the Welsh Definite Article. *Linguistics* 44(4):781–816.

Hayes, Bruce, Péter Siptár, Kie Zuraw & Zsuzsa Londe. 2009. Natural and Unnatural Constraints in Hungarian Vowel Harmony. *Language* 85(4):822–863.

Hayes, Bruce & Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39(3):379–440.

Kim, Eun-Sook. 2003. Theoretical issues in Nuu-chah-nulth Phonology and Morphology (British Columbia). PhD Dissertation, University of British Columbia.

Kurisu, Kazutaka. 2001. The Phonology of Morpheme Realization. PhD Dissertation, University of California, Santa Cruz.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic Grammar – A Formal Multi-Level Connectionist Theory of Linguistic Well-Formedness: Theoretical Foundation. ICS Technical Report 90-5, University of Colorado, Boulder, CO.

Lionnet, Florian. 2017. A Theory of Subfeatural Representations: The Case of Rounding Harmony in Laal. *Phonology* 34(3):523–564.

Mascaró, Joan. 2007. External Allomorphy and Lexical Representation. *Linguistic Inquiry* 38(4):715–735.

McCollum, Adam. 2019. Gradient Morphophonology: Evidence from Uyghur Vowel Harmony. In Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose & Matt Zaslansky (eds.), *Proceedings of the Annual Meeting on Phonology*, Washington, D.C.: LSA.

Pater, Joe. 2009. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Stephen Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*, 123–154. London: Equinox. https://rucore.libraries.rutgers.edu/rutgers-lib/41017/.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker. 2010. Harmonic Grammar with linear Programming: From Linear Systems to Linguistic Typology. *Phonology* 27(1):77–117.

Rosen, Eric. 2016. Predicting the Unpredictable: Capturing the apparent Semi-Regularity of Rendaku Voicing in Japanese through Harmonic Grammar. In Emily Clem, Virginia Dawson, Alice Shen, Amalia Horan Skilton, Geoff Bacon, Andrew Cheng & Erik Hans Maier (eds.), *Proceedings of BLS 42*, 235–249. Berkeley: Berkeley Linguistic Society.

Smolensky, Paul & Matthew Goldrick. 2016. Gradient Symbolic Representations in Grammar: The Case of French Liaison. Ms., Johns Hopkins University.

van Oostendorp, Marc. 2007. Derived Environment Effects and Consistency of Exponence. In Sylvia Blaho, Patrik Bye & Martin Krämer (eds.), *Freedom of Analysis?*, 123–148. Berlin / New York: Mouton de Gruyter.

Zimmermann, Eva. 2018a. Being Exceptional is Being Weak: Tonal Exceptions in San Miguel el Grande Mixtec. In Gillian Gallagher, Maria Gouskova & Sora Yin (eds.), *Proceedings of the 2017 Annual Meeting on Phonology*, Washington, DC: LSA.

———. 2018b. Gradient Symbolic Representations in the Output: A Case Study from Moses Columbian Salishan Stress. In Sherry Hucklebridge & Max Nelson (eds.), *Proceedings of NELS 48*, 275–284. Amherst, MA: GLSA.

———. 2019. Gradient Symbolic Representations and the Typology of Ghost Segments. In Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose & Matt Zaslansky (eds.), *Proceedings of the 2018 Annual Meeting on Phonology*, 1–12. Washington, DC: LSA.

Zsiga, Elizabeth C. 1997. Features, Gestures, and Igbo Vowels: An Approach to the Phonology-Phonetics Interface. *Language* 73(2):227–274.