Gradience in Phonology:
The Argument from Exceptions

PΦF, Princeton

April 06, 2019

Eva Zimmermann

UNIVERSITÄT
LEIPZIG

Main Claim

- The assumption of Gradient Symbolic Representations (Smolensky and Goldrick, 2016; Rosen, 2016) that phonological elements can have different degrees of activation allows a unified explanation for the typology of **phonological exceptions**.
- Predictions about exceptionality patterns:
  - exceptional elements can be **exceptional for multiple processes**
  - different **degrees of exceptionality**
  - **implicational relations** between exceptionality classes within a language

Exceptions: Toy Example

A general phonological rule in Lg1: Nasalization

pok–el → pokel
pon–el → ponẽl (V → Ṽ /[+nasal]__)

1. Exceptional non-undergoer

(for nasalization)

pon– et → ponet, *ponẽt

2. Exceptional non-trigger

(for nasalization)

ton –el → tonel, *tonẽl

3. Exceptional undergoer

(for backness harmony)

pok– ek → pokok, *pokek

4. Exceptional trigger

(for nasalization)

tok –el → tokẽl, *tokel

Pattern 1: A Morpheme is Exceptional for More than one Process

A general phonological rule in Lg2: Nasalization

pok−el → pokel
pon−el → ponẽl (V → Ṽ /[+nasal]__)

1. Exceptional non-undergoer

(for nasalization)

pon− et → ponet, *ponẽt

2. Exceptional non-trigger

(for nasalization)

ton −el → tonel, *tonẽl

3. Exceptional undergoer

(for backness harmony)

pok− et → pokot, *poket

4. Exceptional trigger

(for nasalization)

tok −el → tokẽl, *tokel

Pattern 2: Different Degrees of Exceptionality

Lg3 without backness harmony

pok–el → pokel

Exceptional undergoer I

(for backness harmony)

pok–et → pokot, *poket
put–et → putot, *putet

Exceptional undergoer II

(for backness harmony, parasitic on height)

pok–em → pokom, *pokem
put–em → putem, *putom

Plan

1. Gradient Symbolic Representations: Assumptions
2. Exceptionality for More than one Process
3. Different Degrees of Exceptionality
4. Implicational Relations between Exceptionality Classes
5. Discussion

Gradient Symbolic Representations: Assumptions

Gradient Symbolic Representation (=GSR)

- All linguistic symbols have **activity** that can **gradiently** differ with 1=fully active. (Smolensky and Goldrick, 2016; Rosen, 2016)
- Any change in activity is a faithfulness violation – different activities result in **gradient violations of faithfulness**.
- Elements can be weakly active in the output and thus violate **markedness gradiently**. (Zimmermann, 2017*a*,*b*; Faust and Smolensky, 2017; Jang, 2019; Walker, 2019)
- Grammatical computation modeled inside **Harmonic Grammar** where constraints are weighted. (Legendre et al., 1990; Potts et al., 2010)

GSR: Gradient Constraint Violations

(Cf. Walker (2019) for potential problems and scaling factors as an alternative)

- Weakly active segments:
  - they are **easier to delete** than ‘normal’ segments (=MaxS violated to a lesser degree)
  - it is **costly to realize** them (=activity inserted (1-a) or weak activity in the output (1-b+c))
  - they **violate/satisfy** markedness constraints to a lesser degree

(1) Gradient Activity=gradient constraint violations

$b_1a_1t_1$-$p_{0.5}$		Full!	MaxS	DepS	*CC		
		10	10	10	10		
a.	$b_1a_1t_1p_1$			**-0.5**	**-1**	-15	Only fully active S
b.	$b_1a_1t_1p_{0.5}$	**-0.5**			**-0.75**	-12.5	Faithful realization of weak S
c.	$b_1a_1p_{0.5}$	-0.5	**-1**			-15	Deletion of fully active S
☞ d.	$b_1a_1t_1$		**-0.5**			-5	Deletion of weakly active S

(2) Full!: Assign violation 1-X for every output element with activity X.

Exceptionality for More than one Process

Background: Tones in San Pedro Molinos (=MOL)

- all the data in the following comes from Hunter and Pike (1969)
  variety closely related to San Miguel el Grande Mixtec (Cf. Pike (1944); Mak (1950); Hollenbach (2003); McKendry (2013); theoretical accounts in Goldsmith (1990); Tranel (1995); Zimmermann (2018*a*))
- three level tones high (H; á), mid (M; ā), and low (L; à)

(3) *Tonal contrasts in MOL (Hunter and Pike, 1969, 27)*

tātá-są́	tūtą̄-są́	tūtù-są́
‘my father’	‘my firewood’	‘my paper’
ʔùù ríkī	ʔùù kītī	ʔùù híī
‘two woodpeckers’	‘two animals’	‘two fists’

Process 1: H-Perturbation

- some morphemes trigger an **additional H** that overwrites underlying M or L of the initial TBU of a following morpheme

  (the ‘perturbing’ morphemes found in basically all Otomanguean languages (Dürr, 1987; Pike, 1944; Mak, 1950; Hollenbach, 2003; McKendry, 2013))

(4) H-overwriting

$$\mathbf{XX}^{\mathbf{H}}\ \mathrm{XX} \rightarrow \mathrm{XX}\ \mathbf{H}\mathrm{X}$$

Process 1: H-Perturbation

(5) (Hunter and Pike, 1969, 35-36)

	M1	M2	Surface	Tones
Non-perturbing morphemes				
a.	ʔùʃì ‘ten’	rīŋkī ‘mouse’	ʔùʃì rīŋkī ‘ten mice’	LL MM→LL MM
b.	ʔḭ̄ḭ̄ ‘one’	**sùʧī**$^{\mathbf{H}}$ ‘child’	ʔḭ̄ḭ̄ sùʧī ‘one child’	MM+LM^{H}→MM LM
Perturbing morphemes				
c.	**kṵ̀ṵ̀**$^{\mathbf{H}}$ ‘four’	ʧìká ‘baskets’	kṵ̀ṵ̀ ʧ**í**ká ‘four baskets’	LL^{H} LH→LL **H**H
d.	**ʒāʔā**$^{\mathbf{H}}$ ‘chiles’	ʒìʧí ‘dry’	ʒāʔā ʒ**í**ʧí ‘dry chiles’	MM^{H} LH→MM **H**H
e.	**síví**$^{\mathbf{H}}$ ‘name’	tèē ‘man’	síví t**é**ē ‘name of the man’	HH^{H} LM→HH **H**M
f.	**kītī**$^{\mathbf{H}}$ ‘animal’	kūù ‘to die’	kītī k**ú**ù ‘the animal will die’	MM^{H} ML→MM **H**L

Process 2: H-Spreading after Perturbation

- if a perturbing morpheme precedes a morpheme that ends in an M-toned TBU and is also perturbing, both TBU's of this morpheme become high

(6) H-overwriting and spreading

$$\mathbf{XX^{H}\ XM^{H}} \rightarrow \mathrm{XX}\ \mathbf{HH}$$

Process 2: H-Spreading after Perturbation

(7) (Hunter and Pike, 1969, 35-36)

	M1	M2	Surface	Tones
H-overwriting and spreading				
a.	**síví**^{H}	**sùtʃī**^{H}	síví s**ú**tʃ**í**	HH^{H}+LM^{H}→HH **HH**
	‘name’	‘child’	‘name of the child’	
b.	**síví**^{H}	**kītī**^{H}	síví k**í**t**í**	HH^{H}+MM^{H}→HH **HH**
	‘name’	‘animal’	‘name of the animal’	
c.	**kītī**^{H}	**kāā**^{H}	kītī k**áá**	MM^{H}+MM^{H}→MM **HH**
	‘animal’	‘to eat’	‘the animal will eat’	
No spreading if M2 is not M-final				
d.	**kṵ̀ṵ̀**^{H}	**ʒòò**^{H}	kṵ̀ṵ̀ ʒ**ó**ò	LL^{H}+LL^{H}→LL **H**L
	‘four’	‘mont(H)’	‘four months’	
No spreading if M2 has no floating H				
e.	**síví**^{H}	tèē	síví t**é**ē	HH^{H}+LM→HH **H**M
	‘name’	‘man’	‘name of the man’	

Optionally Perturbing Morphemes as Exceptions

- there are three classes of morphemes in MOL:

  1. non-perturbing ones: XX
  2. perturbing ones: **XX$^{\mathbf{H}}$**
     - trigger H-perturbation
     - trigger H-spreading if they end in an M
  3. **optionally perturbing** ones: **XX$^{\mathbf{(H)}}$**
     - only optionally trigger H-perturbation
     - never trigger H-spreading if they end in an M

  → not optional variation between behaving as morpheme type 1 and 2 but mixture of properties

Optionally Perturbing Morphemes: 1. Optional H-Perturbation

(8) (Hunter and Pike, 1969, 35-36)

	M1	M2	Surface	Tones
a.	**hìkī**$^{\mathbf{(H)}}$	tèē	hìkī t**é**ē∼tèē	LM$^{(H)}$+LM→LM **H**M∼LM
	‘fist, paw’	‘man’	‘the man’s fist’	
b.	**hìkī**$^{\mathbf{(H)}}$	ʧì̢ʔī̢	hìkī ʧ**í̢**ʔī̢∼ʧì̢ʔī̢	LM$^{(H)}$+LM→LM **H**M∼LM
	‘fist, paw’	‘skunk’	‘the skunk’s paw’	
c.	**ñùtī**$^{\mathbf{(H)}}$	ʒìʧí	ñùtī ʒ**í**ʧí∼ʒìʧí	LM$^{(H)}$+LH→LM **H**H∼LH
	‘sand’	‘dry’	‘dry sand’	

Optionally Perturbing Morphemes: 2. No Trigger for H-Spreading

(9) (Hunter and Pike, 1969, 36)

	M1	M2	Surface	Tones
Never a trigger…				
a.	**sívíᴴ** ‘name’	**ʧḭ̀ʔḭ̄⁽ᴴ⁾** ‘skunk’	síví ʧ**ḭ́**ʔḭ̄ ‘name of the skunk’	$HH^{H}+LM^{(H)}\rightarrow HH$ **H**M
b.	**hìkī⁽ᴴ⁾** ‘fist, paw’	**ʧḭ̀ʔḭ̄⁽ᴴ⁾** ‘skunk’	hìkī ʧ**ḭ́**ʔḭ̄∼ʧḭ̀ʔḭ̄ ‘the skunk’s paw’	$LM^{(H)}+LM^{(H)}\rightarrow LM$ **H**M∼LM
…but always an undergoer (if realized)				
c.	**ʧḭ̀ʔḭ̄⁽ᴴ⁾** ‘skunk’	**kāāᴴ** ‘to eat’	ʧḭ̀ʔḭ̄ k**áá**∼kāā ‘the skunk will eat (it)’	$LM^{(H)}+MM^{H}\rightarrow LM$ **HH**∼MM
d.	**hìkī⁽ᴴ⁾** ‘fist’	**sùʧīᴴ** ‘child’	hìkī s**ú**ʧ**í**∼sùʧī ‘the child’s fist’	$LM^{(H)}+LM^{H}\rightarrow LM$ **HH**∼LM

GSR Account: Representational Assumption

- Some morphemes in MOL end in an **unassociated (=floating) H-tone**
- The floating H of some morphemes is **fully active**: H_1
- The floating H of other morphemes is **partially active**: $H_{0.4}$
    - the weakly active $H_{0.4}$ is not a bad enough problem for *Float and is not always associated (=optionality modeled with MaxEnt (Johnson, 2002; Goldwater and Johnson, 2003; Wilson, 2006) where well-formedness is interpreted as probability; calculated with (Hayes, 2009))
    - the weakly active $H_{0.4}$ is not a bad enough problem for the markedness constraint *[MH] triggering H-spreading

Constraints

(10) a. *FLOAT: Assign X violation for every tone T_1 that is not associated to a TBU where X is the activity of T_1.

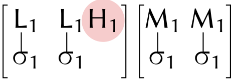
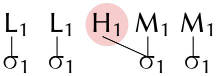
b. MAXT: Assign violation X for any tonal activity X in the input that is not present in the output.

c. *CONT: Assign X violation for every TBU_1 associated to tones T_2 and T_3 where X is the shared activity of TBU_1, T_2, and T_3.

d. SPEC: Assign 1-X violations for every TBU τ_1 where X is the activity of tone(s) associated to τ_1.

H-Perturbation: Realization of H$_1$

(11)

	Max$_H$ 100	*Cont 100	*Float 71	MaxT 24	Spec 8	
a. 			-1			-71
b. 	-1			-1		-124
c. 		-1				-100
☞ d. 				-1		-24

*Float ≫ MaxT

H-Perturbation: Optional Realization of $H_{0.4}$

(12)

$\left[\begin{matrix}L_1 & MH_{0.4}\\ \vert & \vert \\ \sigma_1 & \sigma_1\end{matrix}\right]\left[\begin{matrix}L_1 & M_1\\ \vert & \vert \\ \sigma_1 & \sigma_1\end{matrix}\right]$	Max$_{\text{H}}$ 100	*Cont 100	*Float 71	MaxT 24	Spec 7	
☞ a. $\begin{matrix}L_1 & M & H_{0.4} & L_1 & M_1\\ \vert & \vert & & \vert & \vert \\ \sigma_1 & \sigma_1 & & \sigma_1 & \sigma_1\end{matrix}$			-0.4			-28.4
☞ b. $\begin{matrix}L_1 & M_1 & H_{0.4} & M_1\\ \vert & \vert & \vert & \vert \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1\end{matrix}$				-1	-0.6	-28.2

0.4×*Float ∼ MaxT + 0.6×Spec

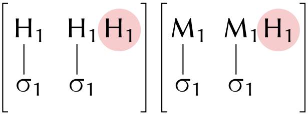
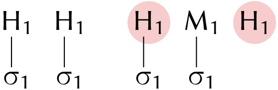
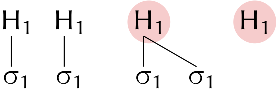
H-Spreading is Avoidance of a Marked Tone Sequence

- triggered by a markedness constraint against sequences of MH-tones inside a morpheme (and only spreading of floating H is a possible repair)

(13) *[MH]: Assign X violation for every morpheme-internal sequence of M_1 and H_2 where X is the shared activity of M_1 and H_2.

H-Spreading Triggered by H_1

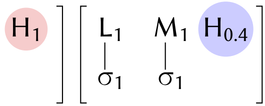
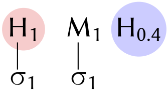
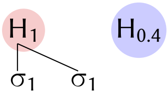
(14)

	Max$_H$ 100	*Float 71	*[MH] 28	MaxT 24	
a. 		-1	-1	-1	-123
☞ b. 		-1		-2	-119

*[MH] ≫ MaxT

No H-Spreading Triggered by Partially Active $H_{0.4}$

(15)

	MAX$_\text{H}$ 100	*FLOAT 71	*[MH] 28	MAXT 24	
☞ a. 		-0.4	-0.7	-1	-72
b. 		-0.4		-2	-76,4

$\text{MAXT} \gg 0.7 \times {}^*[\text{MH}]$

No H-Spreading Triggered by Partially Active $H_{0.4}$

- the assumption of a partially active $H_{0.4}$ predicts the **two exceptional behaviours** from gradient constraint violations

Exceptional optional trigger for H-perturbation

(16) Fully active H_1
*FLOAT ≫ MAXT

(17) Partially active $H_{0.4}$
$0.4\times$*FLOAT ∼ MAXT + $0.6\times$SPEC

Exceptional non-trigger for H-spreading

(18) Fully active H_1
*[MH] ≫ MAXT

(19) Partially active $H_{0.4}$
MAXT ≫ $0.7\times$*[MH]

Different Degrees of Exceptionality

Welsh Ghost Segments 1

(20) Ghost consonant in Welsh (Hannahs and Tallerman, 2006, 798)

a. guda**g** eraill ‘with others’
b. guda gwên ‘with a smile’

Ghost segments: /gudag👻/

Several morphemes surface with an unpredictable consonant only if its appearance avoids a vowel hiatus.

- Those are **appearing ghosts** that only appear to solve a problem.
- Their default state is to **not be there**.

Welsh Ghost segments 2

(21) Welsh definite allomorphy (Hannahs and Tallerman, 2006, 782+783)

a.	**yr** afon	‘the river’	yr (=ər) __ V
b.	**y** llyfr	‘the book’	y (=ə) __ C
c.	o**’r** afon o**’r** llyfr	‘from the river’ ‘from the book’	/’r/ (=r) V__, overriding a.+b.

Ghost segments: /y👻r👻/

A single underlying form /y👻r👻/ and either one of these segments can remain unrealized if it would result in a marked structure (=coda or hiatus).

- Those are **disappearing ghosts** that disappear to avoid a problem.
- Their default state is to **be there**.

Welsh: Ghost segments 1+2 combined

(22) Ghost segments 1+2 combined (Hannahs and Tallerman, 2006, 784)

Underlying	Surface	
gyda**g**👻 **y**👻**r**👻 nod	gyda**'r** nod	'with the aim'

- Why not /gyda**g**👻 **y**👻 nod/ without an additional coda?
  - ➙ because /**g**👻/ is only realized to avoid hiatus, not codas
- Why not /gyda**g**👻 **y**👻**r**👻 nod/ with same additional coda?
  - ➙ because /**g**👻/ is not realized in the default case when markedness is not decisive

Ghosts in Welsh: Summary

(23) *Segments with different behaviour in Welsh*

	default state	non-default state due to	
		*Coda	*Hiat
g👻	not present	no	yes
y👻	present	–	yes
r👻	present	yes	–
n_1	present	no	no

- different thresholds: Is realization more costly or deletion?
- different thresholds: Is *Coda important enough to trigger non-default-state

Welsh Ghost Segments: GSR Account

$$/g_1u_1d_1a_1\mathbf{g}_{\mathbf{0.2}}/ \text{ and } /\mathbf{y}_{\mathbf{0.6}}\mathbf{r}_{\mathbf{0.6}}/$$

- /$\mathbf{y}_{\mathbf{0.6}}$/ and /$\mathbf{r}_{\mathbf{0.6}}$/ are realized unless their realization would create a *Coda or *Hiat violation
- /$\mathbf{g}_{\mathbf{0.2}}$/ is not realized unless it can avoid a *Hiat violation

If only fully active output segments are possible (high-weighted Full!):

(24) */$S_{0.2}$/ is more absent than present*

$g_1u_1d_1a_1\mathbf{g}_{\mathbf{0.2}}$	MaxS 10	DepS 10	
a. $g_1u_1d_1a_1\mathbf{g_1}$		-0.8	-8
☞ b. $g_1u_1d_1a_1$	-0.2		-2

$0.8\times\text{DepS} \gg 0.2\times\text{MaxS}$

(25) */$S_{0.6}$/ is more present than absent*

$\mathbf{y}_{\mathbf{0.6}}\mathbf{r}_{\mathbf{0.6}}$	MaxS 10	DepS 10	
☞ a. $\mathbf{y_1r_1}$		-0.8	-8
b.	-1.2		-12

$0.6\times\text{MaxS} \gg 0.4\times\text{DepS}$

Appearing /$g_{0.2}$/: Realized to Avoid a Vowel Hiatus

(26)

$g_1u_1d_1a_1\mathbf{g_{0.2}}$ $V_1\ldots$	MAXS 10	DEPS 10	*[CC 8	*HIAT 7	*CODA 5	
☞ a. $g_1u_1.d_1a_1.\mathbf{g_{0.2}}V_1$		-0.8				-8
b. $g_1u_1.d_1a_1.V_1$	-0.2			-1		-9

*HIAT + 0.2×MAXS ≫ 0.8×DEPS

Disappearing /$y_{0.6}r_{0.6}$/: /r/ Disappears to Avoid a Coda

(27)

…V_1C_1 **$y_{0.6}r_{0.6}$** C_1V_1…	MAXS 10	DEPS 10	*[CC 8	*HIAT 7	*CODA 5	
a. $V_1.C_1$**y_1r_1**$.C_1V_1$		-0.8			-1	-13
☞ b. $V_1.C_1$**y_1**$.C_1V_1$	-0.6	-0.4				-10
c. $V_1C_1.$**r_1**C_1V_1	-0.6	-0.4	-1		-1	-23
d. $V_1.C_1V_1$	-1.2					-12

*CODA + 0.4×DEPS ≫ 0.6×MAXS

Disappearing /$y_{0.6}r_{0.6}$/: /r/ Disappears to Avoid a Hiatus

(28)

…V_1 **$y_{0.6}r_{0.6}$** V_1…	MAXS 10	DEPS 10	*[CC 8	*HIAT 7	*CODA 5	
a. V_1.**y_1.r_1**V_1		-0.8		-1		-15
b. V_1.**y_1**.V_1	-0.6	-0.4		-2		-24
☞ c. V_1.**r_1**V_1	-0.6	-0.4				-10
d. V_1.V_1	-1.2			-1		-19

*HIAT + 0.4×DEPS $\gg$ 0.6×MAXS

Combination of Appearing and Disappearing Ghosts

(29)

$g_1u_1d_1a_1g_{0.2}\ y_{0.6}r_{0.6}\ C_1V_1\ldots$	RM 100	MaxS 10	DepS 10	*[CC 8	*Hiat 7	*Coda 5	
a. $g_1u._1d_1a_1.g_1y_1r_1.C_1V_1$			-1.6			-1	-21
b. $g_1u._1d_1a_1.y_1r_1.C_1V_1$		-0.2	-0.8		-1	-1	-22
☞ c. $g_1u._1d_1a_1r_1.C_1V_1$		-0.8	-0.4			-1	-17
d. $g_1u._1d_1a_1.g_1y_1.C_1V_1$		-0.6	-1.2				-18

- ➙ vs. (29-d): /$g_{0.2}$/ **never shows its non-default state to avoid codas**
  0.8×DepS ≫ *Coda
- ➙ vs. (29-a): /$g_{0.2}$/ is an appearing ghost and its **default state is thus to not be there**
  0.8×DepS ≫ 0.2×MaxS

Ghosts in Welsh: Summary

(30) *Segments with different behaviour in Welsh*

	default state	non-default state due to	
		*Coda	*Hiat
$g_{0.2}$	not present (24)	no (29)	yes (26)
$y_{0.6}$	present (25)	–	yes (29)
$r_{0.6}$	present (25)	yes (27)	–
n_1	present (34)	no (34)	no (34)

- different thresholds: Is realization more costly or deletion?
- different thresholds: Is *Coda important enough to trigger non-default-state

Implicational Relations between Exceptionality Classes

GSR Prediction: Implicational Relations

(31) Impossible exceptionality pattern (from strength differences alone)

	Triggers/undergoes Process P_1	Triggers/undergoes Process P_2
Normal element	✔	✔
Exceptional element 1	✔	✖
Exceptional element 2	✖	✖
* Exceptional element 3	✖	✔

Implicational Relations: Example MOL^{I}

- two additional exceptional morpheme(s) (classes) 2+4 are possible
- exceptional morpheme class 5 is impossible

(32)

		HP	HS	WA: HP	WA:HS
1.	H_1	✔	✔	*FLOAT ≫ MAXT	*[MH] ≫ MAXT
2.	$H_{0.6}$	✔	(✔)	0.6×*FLOAT ≫ MAXT + 0.4×SPEC	0.6×*[MH] ∼ MAXT
3.	$H_{0.4}$	(✔)	✖	0.4×*FLOAT ∼ MAXT + 0.6×SPEC	MAXT ≫ 0.4×*[MH]
4.	$H_{0.2}$	✖	✖	MAXT + 0.8×SPEC ≫ 0.2×*FLOAT	MAXT ≫ 0.2×*[MH]
*5.	$H_?$	✖	✔	MAXT + (1-?)×SPEC ≫ ?×*FLOAT	?×*[MH] ≫MAXT

HP=trigger for H-perturbation HS=trigger for H-spreading (if ending in M)

✔=yes (✔)=optional ✖=no

Discussion

Further Prediction: True gradience

- no inherent restriction on gradient contrasts within a language
  - 3 types of segments in Welsh:
    $/k_{1.0}/$ - $/r_{0.6}/$ - $/g_{0.2}/$
  - 3 types of association lines in Oku (Trommer and Zimmermann, 2018):
    $/H\!-_{1.0}\!\bullet/$ - $/H\!-_{0.6}\!\bullet/$ - $/H\!-_{0.4}\!\bullet/$
  - 4 (derived) segment types in Levantine Arabic (Trommer, 2018*a*):
    $/i_{0.7}/$ - $/i_{0.6}/$ - $/i_{0.5}/$ - $/i_{0.3}/$
  - 5 types of feet in Moses Columbian Salish (Zimmermann, to appear):
    $/\varphi_{1.0}/$ - $/\varphi_{0.9}/$ - $/\varphi_{0.8}/$ - $/\varphi_{0.6}/$ - $/\varphi_{0.4}/$

- vs. alternatives
  - most accounts based on autosegmental defectivity that only allow a binary distinction into [±defective] (e.g. Hyman, 1985; Noske, 1985; Kenstowicz and Rubach, 1987; Sloan, 1991; Yearley, 1995; Tranel, 1996; Zoll, 1996)
  - accounts that adopt ‘strength’ as a binary division
    (Inkelas, 2015; Vaxman, 2016*a*,*b*; Sande, 2017)

Further Prediction: Surface activity and phonetic interpretation

- phonetic gradience in phonology:
  - subphonemic gradience in word-final devoicing, nasal place assimilation, flapping (Braver, 2013, e.g.)
  - vowel harmony is gradient; gets weaker the farther it spreads (McCollum, 2018)

→ a convincing example would be one where phonetic gradience and exceptional phonological behaviour stemming from underlying weakness coincide

Open Question: The source for strength

- lexical contrast for phonological elements
- lexical contrast for whole morphemes (Faust and Smolensky, 2017)
- **derived in the phonology**:
  - 'Gradient representations can mature or decay across layers' (Trommer, 2018*a*)
  - stress strengthens elements (Faust and Smolensky, 2017; Amato, 2018; Trommer, 2018*a*)
  - floating strength strengthens elements (Amato, 2018)
  - fission is weakening/distribution of activity (Zimmermann, 2019)
  - certain features have an inherent strength and feature change thus implies strength adjustment (Walker, 2019)

Arguments for GSR

1. Embedded in a general **computational architecture for cognition**
   (=Gradient Symbolic Computation Smolensky and Goldrick, 2016)
2. A **unified account** for different exceptional phonological behaviours:
   - liaison consonants in French (Smolensky and Goldrick, 2016)
   - semi-regularity of voicing in Japanese Rendaku (Rosen, 2016)
   - allomorphy in Modern Hebrew (Faust and Smolensky, 2017)
   - lexical accent in Lithuanian (Kushnir, 2017)
   - tone sandhi in Oku (Nformi and Worbs, 2017)
   - tone allomorphy in San Miguel el Grande Mixtec (Zimmermann, 2017*a*,*b*)
   - lexical stress in Moses Columbian Salishan (Zimmermann, to appear)
   - exceptional tone (non)spreading in San Molinos Mixtec (Zimmermann, 2018*b*)
   - interaction of phonological/lexical gemination/lenition in Italian (Amato, 2018)
   - compound stress in Sino-Japanese (Rosen, 2018)
   - stress-syncope interaction in Levantine Arabic (Trommer, 2018*b*)
   - (interacting) ghost segments in Welsh (Zimmermann, 2018*c*)
   - …

Summary

- the assumption of GSR with activity in the output predicts the typology of exceptions from **gradient faithfulness and markedness violations**
- certain elements are predicted to be **exceptional for more than one process**
  - argument against lexically indexed constraints (e.g. Alderete, 2001; Pater, 2010; Finley, 2009))
  - argument against autosegmental defectivity accounts since gradient violations of constraints directly referring to this element are sufficient
- different **grades of exceptionality** are predicted
- **implicational restrictions** between exceptions are predicted

References

Alderete, John (2001), *Morphologically governed accent in Optimality Theory*, Routledge, New York.

Amato, Irene (2018), ‘A gradient view of Raddoppiamento Fonosintattico’, ms., University of Leipzig.

Braver, Aaron (2013), Degrees of incompleteness in neutralization: Paradigm uniformity in a phonetics with weighted constraints, PhD thesis, Rutgers The State University of New Jersey-New Brunswick.

Dürr, Michael (1987), ‘A preliminary reconstruction of the Proto-Mixtec tonal system’, *Indiana* **11**, 19–61.

Faust, Noam and Paul Smolensky (2017), ‘Activity as an alternative to autosegmental association’, talk given at mfm 25, 27th May, 2017.

Finley, Sara (2009), ‘Morphemic harmony as featural correspondence’, *Lingua* **119**, 478–501.

Goldsmith, John (1990), *Autosegmental and Metrical Phonology*, Blackwell, Oxford.

Goldwater, Sharon and Mark Johnson (2003), Learning ot constraint rankings using a maximum entropy model, *in* J.Spenader, A.Eriksson and O.Dahl, eds, ‘Proceedings of the Workshop on Variation within Optimality Theory’, Stockholm University, Stockholm, pp. 111–120.

Hannahs, S. J. and Maggie Tallerman (2006), ‘At the interface: selection of the Welsh definite article’, *Linguistics* **44**, 781–816.

Hayes, Bruce (2009), ‘Manual for maxent grammar tool’, online available at http://linguistics.ucla.edu/people/hayes/MaxentGrammarTool/ManualForMaxentGrammarTool.pdf.

Hollenbach, Barbara (2003), The historical source of an irregular Mixtec tone-sandhi pattern, *in* M. R.Wise, T.Headland and R.Brend, eds, ‘Language and life: essays in memory of Kenneth L. Pike’, SIL International, Dallas, pp. 535–552.

Hunter, Georgia and Eunice Pike (1969), ‘The phonology and tone sandhi of Molinos Mixtec’, *Linguistics* .

Hyman, Larry (1985), *A theory of phonological weight*, Foris Publications, Dordrecht.

Inkelas, Sharon (2015), Confidence scales: A new approach to derived environment effects, *in* Y. E.Hsiao and L.-H.Wee, eds, ‘Capturing Phonological Shades Within and Across Languages’, Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 45–75.

Jang, Hayeun (2019), ‘Emergent phonological gradience from articulatory synergies: simulations of coronal palatalization’, talk, presented at the LSA 2019, New York, January 05, 2019.

Johnson, Mark (2002), Optimality-theoretic lexical functional grammar, *in* S.Stevenson and P.Merlo, eds, ‘The Lexical Basis of Sentence Processing: Formal, Computational and Experimental Issues’, John Benjamins, Amsterdam, pp. 59–73.

Kenstowicz, Michael and Jerzy Rubach (1987), ‘The phonology of syllabic nuclei in Slovak’, *Language* **63**, 463–497.

Kushnir, Yuriy (2017), ‘Accent strength in Lithuanian’, talk, given at the workshop on Strength in Grammar, Leipzig, November 12, 2017.

Legendre, Geraldine, Yoshiro Miyata and Paul Smolensky (1990), ‘Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations’, *Proceedings of the 12th annual conference of the cognitive science society* pp. 388–395.

Mak, Cornelia (1950), ‘A unique tone perturbation in Mixteco’, *International Journal of American Linguistics* **16**, 82–86.

McCollum, Adam (2018), ‘Gradient morphophonology: Evidence from Uyghur vowel harmony’, talk at AMP 2018, San Diego, October 06, 2018.

McKendry, Inga (2013), Tonal Association, Prominence and Prosodic Structure in South-Eastern Nochixtlán Mixtec, PhD thesis, University of Edinburgh.

Nformi, Jude and Sören Worbs (2017), ‘Gradient tones obviate floating features in Oku tone sandhi’, talk at the Workshop on Strength in Grammar, Leipzig, November 10, 2017.

Noske, Roland (1985), Syllabification and syllable changing processes in Yawelmani, *in* H.van der Hulst and N.Smith, eds, 'Advances in Nonlinear Phonology', Foris, pp. 335–361.

Pater, Joe (2010), Morpheme-specific phonology: Constraint indexation and inconsistency resolution, *in* S.Parker, ed., 'Phonological Argumentation: Essays on Evidence and Motivation', Equinox, London, pp. 123–154.

Pike, Kenneth L. (1944), 'Analysis of a Mixteco text', *International Journal of American Linguistics* **10**, 113–138.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt and Michael Becker (2010), 'Harmonic grammar with linear programming: From linear systems to linguistic typology', *Phonology* pp. 77–117.

Rosen, Eric (2016), Predicting the unpredictable: Capturing the apparent semi-regularity of rendaku voicing in Japanese through Harmonic Grammar, *in* E.Clem, V.Dawson, A.Shen, A. H.Skilton, G.Bacon, A.Cheng and E. H.Maier, eds, 'Proceedings of BLS 42', Berkeley Linguistic Society, Berkeley, pp. 235–249.

Rosen, Eric (2018), 'Evidence for gradient input features from Sino-Japanese compound accent', poster, presented at AMP 2018, San Diego, October 06, 2018.

Sande, Hannah (2017), Distributing morphologically conditioned phonology: Three case studies from Guébie, PhD thesis, University of California, Berkeley.

Sloan, Kelly Dawn (1991), Syllables and Templates: Evidence from Southern Sierra Miwok, PhD thesis, MIT.

Smolensky, Paul and Matthew Goldrick (2016), 'Gradient symbolic representations in grammar: The case of French liaison', Ms, Johns Hopkins University and Northwestern University, ROA 1286.

Tranel, Bernard (1995), 'Rules vs. constraints: a case study', ROA 72.

Tranel, Bernard (1996), Exceptionality in optimality theory and final consonants in French, *in* K.Zagona, ed., ‘Grammatical Theory and Romance Languages: Selected papers from the 25th Linguistic Symposium on Romance Languages (LSRL XXV)’, John Benjamins, Amsterdam, pp. 275–291.

Trommer, Jochen (2018*a*), ‘The layered phonology of Levantine Arabic syncope’, talk at the Workshop on Cyclic Optimization, Leipzig, May 18, 2018.

Trommer, Jochen (2018*b*), ‘The layered phonology of Levantine Arabic syncope’, poster, presented at AMP 2018, San Diego, October 07, 2018.

Trommer, Jochen and Eva Zimmermann (2018), ‘The strength and weakness of tone: A new account to tonal exceptions and tone representations’, invited talk, given at the Phorum, UC Berkeley, March 19, 2018.

Vaxman, Alexandre (2016*a*), Diacritic weight in the extended accent first theory, *in* ‘University of Pennsylvania Working Papers in Linguistics’, University of Pennsylvania.

Vaxman, Alexandre (2016*b*), How to Beat without Feet: Weight Scales and Parameter Dependencies in the Computation of Word Accent, PhD thesis, University of Connecticut.

Walker, Rachel (2019), ‘Gradient feature activation and the special status of coronals’, talks, presented at PΦF 2019, April 05, 2019.

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

Yearley, Jennifer (1995), Jer vowels in Russian, *in* J.Beckman, L.Walsh Dickey and S.Urbanczyk, eds, ‘Papers in Optimality Theory’, GLSA Publications, Amherst, Mass., pp. 533–571.

Zimmermann, Eva (2017*a*), ‘Being exceptional is being weak: tonal exceptions in San Miguel el Grande Mixtec’, poster, presented at AMP 2017, New York, September 16, 2017.

Zimmermann, Eva (2017*b*), ‘Gradient symbols and gradient markedness: a case study from Mixtec tones’, talk, given at the 25th mfm, 27th May, 2017.

Zimmermann, Eva (2018*a*), Being exceptional is being weak: Tonal exceptions in San Miguel el Grande Mixtec, *in* G.Gallagher, M.Gouskova and S. H.Yin, eds, 'Proceedings of AMP 2017', LSA, http://dx.doi.org/10.3765/amp.

Zimmermann, Eva (2018*b*), 'Exceptional non-triggers are weak: The case of Molinos Mixtec', talk at OCP 15, January 13, 2018.

Zimmermann, Eva (2018*c*), 'Gradient symbolic representations and the typology of ghost segments: An argument from gradient markedness', talk, given at AMP 2018, San Diego, October 06, 2018.

Zimmermann, Eva (2019), 'Faded copies: Reduplication as sharing of activity', talk, to be given at OCP 16.

Zimmermann, Eva (to appear), 'Gradient symbolic representations in the output: A case study from Moses Columbian Salishan stress', *Proceedings of NELS 48* .

Zoll, Cheryl (1996), Parsing below the segment in a constraint-based framework, PhD thesis, UC Berkeley.

`Eva.Zimmermann@uni-leipzig.de`

Welsh: Ghost segments 1+2 combined

(33) Underlying: /gyda**g**👻 **y**👻**r**👻 nod/ (Hannahs and Tallerman, 2006, 784)

		ghost deleted	ghost realized	marked
* Option 1:	gyda**g** **y** nod	☹	☹ ☹	☺
☞ Option 2:	gyda**'r** nod	☹ ☹	☹	☹

'with the aim'

Realization of /r/ takes precedence over the other ghost segments

- one of the reasons Hannahs and Tallerman (2006) reject a phonological account of the definite allomorphy
- → follows in an account based on gradient activity where segment can have different default states: /g/'s **default state is not to be there**

Welsh: Markedness and Non-Ghosts in Welsh

- non-ghost segments are neither deleted nor inserted to avoid *Hiat and/or *Coda problems

(34)

$\ldots V_1\ a_1f_1o_1n_1\ C_1V_1\ldots$	MaxS	DepS	*[CC	*Hiat	*Coda	
	10	10	8	7	5	
☞ a. $V_1.a_1.f_1o_1n_1.C_1V_1$				-1	-1	-12
b. $V_1.a_1.f_1o_1.C_1V_1$	-1			-1		-17
c. $V_1.\text{ʔ}_1a_1.f_1o_1n_1.C_1V_1$		-1			-1	-15
d. $V_1.\text{ʔ}_1a_1.f_1o_1.C_1V_1$	-1	-1				-20

MaxS $\gg$ *Coda/*Hiat
DepS $\gg$ *Coda/*Hiat

MOL: Perturbing Morphemes: Summary

- the optionally perturbing morphemes
  - only optionally trigger H-Perturbation
  - never trigger H-Spreading

M2 M1	XX	**XM$^{(H)}$**	**XM^{H}**	
XX	no change	no change	no change	
XX$^{(H)}$	no change or H-OW	no change or H-OW	no change or H-OW+Spr	⇒ **Sometimes H-OW trigger**
XX^{H}	H-OW	H-OW	H-OW+Spr	⇒ **Always H-OW trigger**
		⇓ **Never H-Spr trigger**	⇓ **Always H-Spr trigger**	

MOL: No repair possible for *[MH] without a floating H

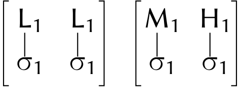
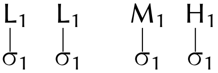
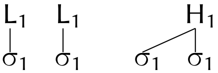
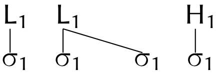
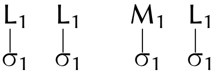
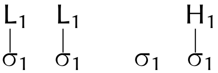
- simply deleting a tone is excluded by SPECIFY (=SPEC)
- deleting a tone and inserting one is excluded by DEPT
- spreading an underlying tone of the same morpheme is excluded by ALTERNATION
- spreading an underlyingly associated tone of a preceding morpheme is excluded by $^{*}\text{LONG}_{\text{MBOUND}}$ (35)

(35) a. $^{*}\text{LGT}_{M}$: Assign X violations for every tone T_1 that is associated to two TBU's τ_2 and τ_3 of different morphological affiliations where X is the shared activity of T_1, τ_2, and τ_3.

b. SPEC: Assign 1-X violations for every TBU τ_1 where X is the activity of tone(s) associated to τ_1.

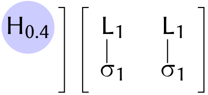
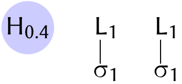
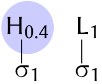
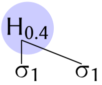
MOL: No repair possible for *[MH] without a floating H

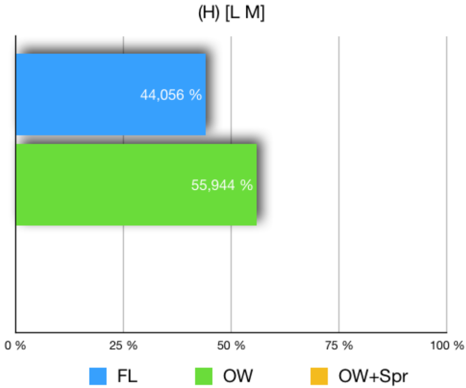
(36)

		Alt 100	*LgT$_M$ 100	DepT 100	*[MH] 28	MaxT 24	Spec 7	
☞ a.					-1			-28
b.		-1				-1		-124
c.			-1			-1		-124
d.				-1		-1		-124
e.						-1	-1	-31

MOL: H-Perturbation: Optional Realization of $H_{0.4}$: MaxEnt

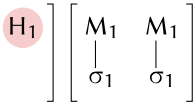
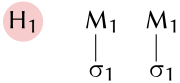
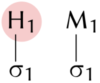
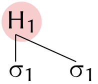
(37)

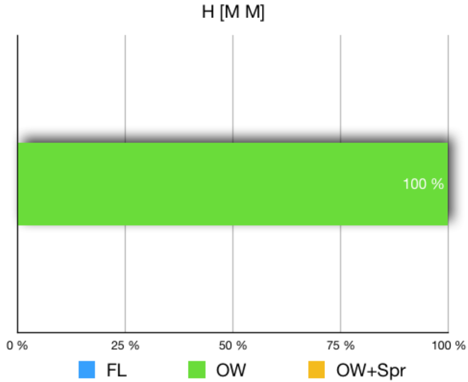
	*H*	Probability
☞ a. 	-28,4	0,4406
☞ b. 	-28,16	0,5594
c. 	-34,5	3,29E-13



MOL: Fully active H_1 is realized: Maxent probabilities

(38)

		*H*	Probability
a.		-71,0	4,20E-21
☞ b.		-24,08	0,9999
c.		-48,16	3,49E-11



GSR Prediction: Implicational Relations, Toy Example

- Given a Lg where unexceptional S never undergoes process P_1 or P_2 to avoid the markedness violations M_1 or M_2 respectively:
- (and only gradient faithfulness is relevant, not gradient markedness)

(39)

	P_1	P_2		
S_1	✖	✖	Faith $\gg M_1$	Faith $\gg M_2$
$S_{0.6}$	✔	✖	$M_1 \gg 0.6\times$Faith	$0.6\times$Faith $\gg M_2$
$S_{0.4}$	✔	✔	$M_1 \gg 0.4\times$Faith	$M_2 \gg 0.4\times$Faith
* $S_{???}$	✖	✔	??? x Faith $\gg M_1$	$M_2 \gg$??? x Faith