A typology of OCP-resolutions across morphological domains: An argument for Harmonic Layer Theory

Eva Zimmermann
MST-Project

24. Mai 2024

UNIVERSITÄT
LEIPZIG

Main Claim

- Harmonic Layer Theory where phonological elements can get incrementally stronger/weaker at every optimization cycle predicts different at morphosyntactic layers behaviour from **a single phonological grammar**
- In contrast to multi-grammar alternatives, the theory is **restricted by monotonicity** and excludes unattested XYX patterns.
- Lexical exceptions fall out from the same general mechanism (=activity differences) than layer asymmetries

Harmonic Layer Theory and Shona

Data: Different morphosyntactic domains in Shona

Shona

- Zezuru dialect, spoken in Zimbabwe
  (Atlantic-Congo –> East Bantu –> Southern Bantu)
- sources for data and generalizations: Myers (1997) and Myers (1986)
- TBU's = syllables
- high-toned (=V́) or low-toned (=V)
- L-tones are taken to be (underlyingly) absent/inserted later

Relevant processes in Shona: OCP contexts

(1) Context 1: A singly-linked first H

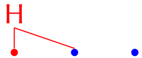
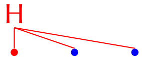
H H	OCP$_{\text{TBU}}$	MaxH	‖Unif‖
a. H H –	-1		
b. H Del		-1	
c. H Fus			-1

(2) Context 2: A multiply-linked first H

H H	OCP$_{\text{TBU}}$	MaxH	‖Unif‖
a. H H Ret			

Relevant processes in Shona: Spreading contexts

(3) Spreading: Long or short

	Spec	‖*LongH$_3$‖	‖*LongH‖
a. 	-2		
b.  H1S	-1		-1
c.  H2S		-1	-1

Different morpho-syntactic domains=different phonologies

(4) Structure of the Shona verb

Word			
[Macrostem]	**[Macrostem]**		
Clitics (copula, Assoc., remaining inflection)	prefixes– (object, some subject, some Tns)	root	–suffixes
ku	mú̱-	té̱ng	-és-ér-a
INF	OBJ	buy	CAUS-applied-FV

(5) Different behaviour at different layers: Overview

	OCP context	H-spread
1 (Macro)Stem	Ret > Fus	H2S
2 Word	Ret > Del	H1S
3 Phrase	Ret > –	H1S

Examples: Spreading asymmetries

(underlying H-tones are notated with v̲, surface H-tones with v́)

(6)

a. H2S at 1, triggered by Obj

[tí̲-táríse]

1PL/SUBJ-look-FV

'we would look' (M7:870)

b. H1S at 2; triggered by copula

[í̲]-[sádza]

COP-porridge

'(it) is porridge' (M7:860)

c. H2S at 1 + H1S at 3

[ku]-[té̲ng-és-á] [sádza]

INF-buy-CAUS-FV porridge

'to sell porridge' (M7:862)

d. H1S at 2 blocked by OCP

[í̲]-[badzá̲]

COP-hoe

'(it) is a hoe' (M7:860)

OCP asymmetries

(7)

a. Fus at 1 between Obj+root

[ku]-[mú-téng-és-ér-a]

INF-OBJ-buy-CAUS-applied-FV

'to sell him/her' (M7:869)

b. Del at 2

[í]-[banga]

COP-knife

'(it) is a knife' (M7:856)

OCP asymmetries

(8)

Ret at 2, fed by Fus at 1

[á-cha]-[téng-á]

3SG-FUT=buy-FV

's/he will buy' (M7:864)

(9)

H H a cha	H teng a	underlying representations
Fus H a cha	H2S H teng a	1 Two macrostems
Ret H H a cha teng a		2 One Word

OCP asymmetries

(10)

a. Del at 2, fed by Fus at 1

[há]-[ti-teng-es-e]

HORT-1PL/SUBJ-buy-CAUS-FV

'let us sell' (M7:870)

b. – if Ret impossible at 3

[badzá] [gúrú]

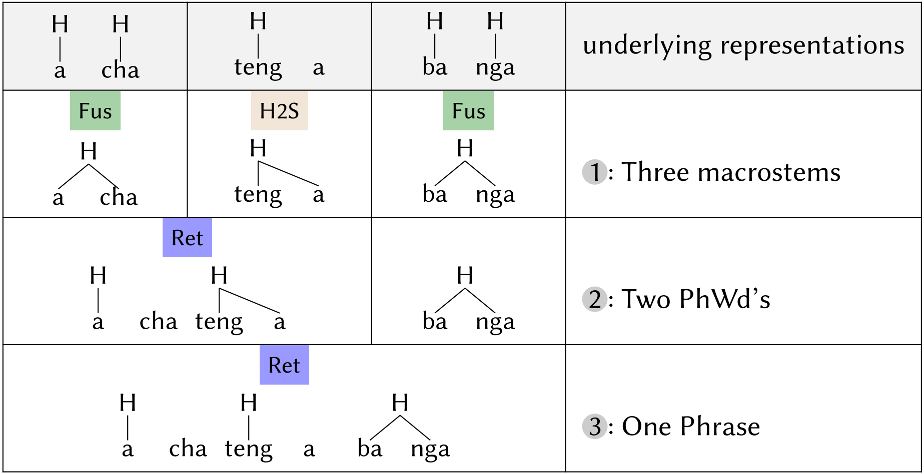
hoe big

'big hoe' (M7:874, FN.21)

(11)

H [ha]	H H [ti teng es e]	underlying representations
H [ha]	Fus + H2S H [ti teng es e]	1 Two macrostems
Del H [ha] [ti teng es e]		2 One word

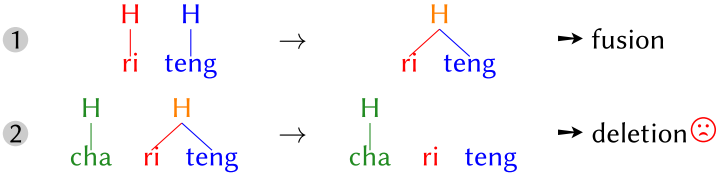
A sample derivation



An exception: Object /rí̲-/

(12) [ku]-[**rí̲**-té̲ng-és-á]
INF-OBJ/5-buy-CAUS-FV
'to sell it' (Myers, 1986, 197)

(13) Our expectation



(14) [ndi-chá̲]-[**ri**-té̲ng-á]
1SG-FUT-OBJ-buy-FV
'I will buy it' (Myers, 1986, 211)

- /rí̲/ surfaces with its H in 1 (with or without a following H)
- but it exceptionally is the only undergoer of H-deletion at 2

➙ /rí̲-/ fails to undergo fusion at 1

Shona: Empirical summary

(15)

	OCP context				H-spread	
1 (Macro)Stem	Ret	>	Fus	*	H2S	*exceptional non-undergoer /rí/
2 Word	Ret	>	Del		H1S	
3 Phrase	Ret	>	–		H1S	

HLT: The assumptions

Harmonic Layer Theory: Defining assumptions

1. a **single constraint ranking** for a grammar
2. cyclic evaluation at **three layers** (=stem, word, phrase)
3. phonological elements can have **gradient activity** that results in gradient constraint violations
4. activity can predictably **decay** at every optimization

Background: Gradient Symbolic Representations (=GSR)

- all linguistic symbols have **activity** that can **gradiently** differ (Smolensky and Goldrick, 2016; Rosen, 2016, 2019*a*,*b*; Zimmermann, 2018*a*, 2019; Kushnir, 2019, e.g.)
- different activities result in **gradient constraint violations** of both faithfulness and markedness constraints (Zimmermann, 2021, to appear; 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)

cite something

GSR and constraint violations

- gradient constraints are **violated/satisfied relative to the activity** of the relevant elements

(16) a. $\textsc{Max}_H$: Assign -x violation for every H_x in the input without an output correspondent.

b. $\textsc{Max}_L$: Assign -x violation for every L_x in the input without an output correspondent.

c. *H: Assign -x violation for every H_x.

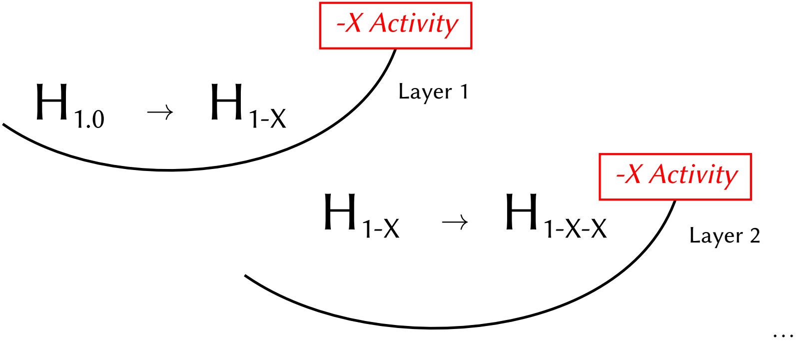
d. *L: Assign -x violation for every L_x.

(17) Gradient activity=gradient constraint violations

$H_1\ L_{0.8}$	$\textsc{Max}_H$ 20	$\textsc{Max}_L$ 20	*H 10	*L 10	
☞ a. $H_1\ L_{0.8}$			-1.0	-0.8	-18
b. $L_{0.8}$	-1.0			-0.8	-28
c. H_1		-0.8	-1.0		-26

HLT: Predictable loss of activity at every layer

- constraint interaction can ensure that all instances of a certain element (e.g. H) **loose a fixed amount of activity** at every optimization cycle



Harmonic Layer Theory: Assumptions open to debate

1. **constraint cloning**: Every faithfulness and markedness constraint exists in a gradient and ‖categorical version‖

   (18) a. MAX$_H$: Assign -x violation for every H_x in the input without an output correspondent.

   b. ‖MAX$_H$‖: Assign -1 violation for every H in the input without an output correspondent.

   c. *H: Assign -x violation for every H_x.

   d. ‖*H‖: Assign -1 violation for every H.

2. **Egalitarian** optimization: Every morpheme is optimized at every layer

3. the default underlying activity (assumed by the learner) is 1

4. activity can only decay across layers

A HLT account of Shona

HLT account of Shona

1. tones get weaker at each layer
2. gradient markedness:
   - spreading becomes a worse repair the weaker the tone becomes
   - the OCP problem gets less severe the weaker the tone
3. gradient faithfulness:
   - deletion as a repair gets cheaper the weaker the tone
4. lexical activity differences: object /rí̱/ is weak and hence doesn't demand an OCP-repair

Activity loss at every optimization: Logic

- constraint interaction ensures that all **H's decay** by 0.2 at every layer

(19) Predictable decay by 0.2

	INPUT	OUTPUT
1	H_1	$H_{0.8}$
2	$H_{0.8}$	$H_{0.6}$
3	$H_{0.6}$	$H_{0.4}$

(20) a. *H: Assign -x violation for every TBU associated to H_x.

b. IDACT: Assign -x violation for every input tone H_a corresponding to output tone H_b where a-b=x and x is > 0.

c. $\text{IDACT}_{\leq 0.2}$: Assign -x violation for every input tone H_a corresponding to output tone H_b where a-b=x and x is > 0.2.

Activity loss at every optimization: Tableaux

1

	$H_{1.0}$	IDACT$_{\leq 0.2}$ w=∞	MAX$_H$ w=20	*H w=10	IDACT w=1	$\mathcal{H}$
	a. $H_{1.0}$			-1.0		-10.0
	b. Ø		-1.0			-20.0
	c. $H_{0.5}$	-0.5		-0.5	-0.5	∞
☞	d. $H_{0.8}$			-0.8	-0.2	-8.2

2

	$H_{0.8}$	IDACT$_{\leq 0.2}$ w=∞	MAX$_H$ w=16	*H w=15	IDACT w=1	$\mathcal{H}$
	a. $H_{0.8}$			-0.8		-8.0
☞	b. $H_{0.6}$			-0.6	-0.2	-6.2

3

	$H_{0.6}$	IDACT$_{\leq 0.2}$ w=∞	MAX$_H$ w=16	*H w=15	IDACT w=1	$\mathcal{H}$
	a. $H_{0.6}$			-0.6		-6.0
☞	b. $H_{0.4}$			-0.4	-0.2	-4.2

Different behaviour for spreading: In a nutshell

- spreading applies to provide otherwise toneless TBU's with a tone
- but this strategy gets less helpful, the weaker the tone is

(21) a. SPEC
Assign -(1-x) violation for every TBU associated with tone T_x (and no tone=T_0).

b. ||*LONGH||
Assign -1 violation for every tone that is associated to more than one TBU.

c. ||*LONGH$_3$||
Assign -1 violation for every tone that is associated to more than two TBU's.

Tableaux: H2S at 1 but H1S at 2

1 H_1		SPEC 8	$\|$*LONGH$_3\|$ 6	$\|$*LONGH$\|$ 0	$\mathcal{H}$
a. $H_{0.8}$	–	-2.2			-17.6
b. $H_{0.8}$	H1S	-1.4		-1.0	-11.2
☞ c. $H_{0.8}$	H2S	-0.6	-1.0	-1.0	-10.8

2 $H_{0.8}$		SPEC 8	$\|$*LONGH$_3\|$ 6	$\|$*LONGH$\|$ 0	$\mathcal{H}$
a. $H_{0.6}$	–	-2.4			-19.2
☞ b. $H_{0.6}$	H1S	-1.8		-1.0	-14.4
c. $H_{0.6}$	H2S	-1.2	-1.0	-1.0	-15.6

Spreading asymmetry: The weighting arguments

(22)

1	H2S	0.8xSpec	>	$\|$*LongH$_3$$|$
2	H1S	$\|$*LongH$_3$$|$	>	0.6xSpec
3	H1S	$\|$*LongH$_3$$|$	>	0.4xSpec

Different behaviour for OCP problems: In a nutshell

(23) Different behaviour at different layers: Overview

	OCP avoided by:	H-spread
1 (Macro)Stem	Ret > Fus	H2S
2 Word	Ret > Del	H1S
3 Phrase	Ret > –	H1S

- fusion is the default repair but:
  - the weaker the H, the cheaper deletion
  - the weaker the H, the easier it is to tolerate the OCP
- (additional constraint ||DEL2!|| favours deletion of second over deletion of first H)

(24)

a. OCP$_\bullet$: Assign -x violation for every H_x followed by another H if both are associated with adjacent TBU's.

b. MAX$_H$: Assign -x violation for every H_x in the input without an output correspondent.

c. ||UNIF||: Assign -1 violation for every pair of input tones corresponding to the same output tone.

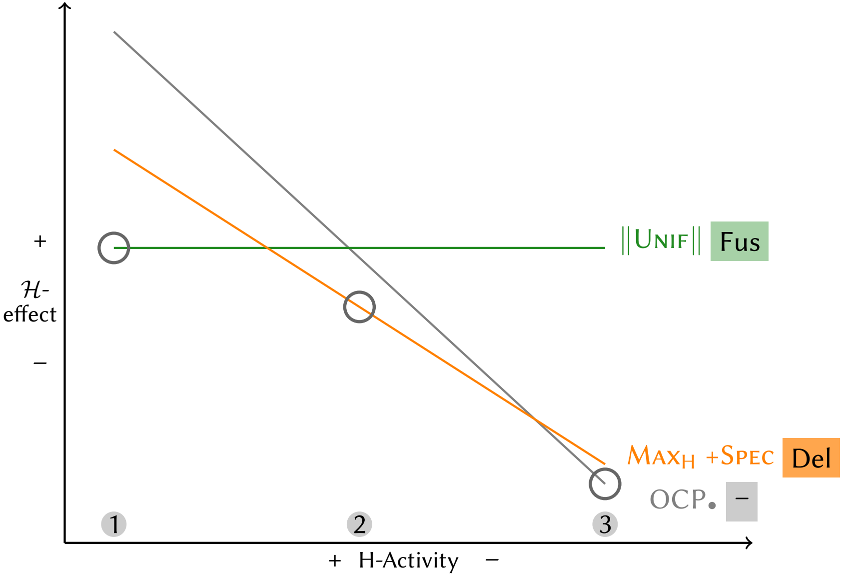
Three layers and the OCP

1	$H_{1.0}$ $H_{1.0}$		OCP• 27	MAX$_H$ 20	‖UNIF‖ 17	*H 10	SPEC 8	$\mathcal{H}$
a.	$H_{0.8}$ $H_{0.8}$	–	-0.8			-1.6	-0.4	-40.8
b.	$H_{0.8}$	Del		-1.0		-0.8	-1.2	-37.6
☞ c.	–$H_{0.8}$–	Fus			-1.0	-1.6	-0.4	-36.2

2	$H_{0.8}$ $H_{0.8}$		OCP• 27	MAX$_H$ 20	‖UNIF‖ 17	*H 10	SPEC 8	$\mathcal{H}$
a.	$H_{0.6}$ $H_{0.6}$	–	-0.6			-1.2	-0.8	-34.6
☞ b.	$H_{0.6}$	Del		-0.8		-0.6	-1.4	-33.2
c.	–$H_{0.6}$–	Fus			-1.0	-1.2	-0.8	-35.4

3	$H_{0.6}$ $H_{0.6}$		OCP• 27	MAX$_H$ 20	‖UNIF‖ 17	*H 10	SPEC 8	$\mathcal{H}$
☞ a.	$H_{0.4}$ $H_{0.4}$	–	-0.4			-0.8	-1.2	-28.4
b.	$H_{0.4}$	Del		-0.6		-0.4	-1.6	-28.8
c.	–$H_{0.4}$–	Fus			-1.0	-0.8	-1.2	-34.6

Shona: HLT thresholds, OCP contexts



Multiply linked H1: Retraction is always cheaper (MStem)

(25) $\|$MAX$_{|}\|$: Assign -1 violation for every input association between elements A and B that is not realized in the output.

1 H_1 H_1	OCP• 27	MAX$_H$ 20	$\|$UNIF$\|$ 17	$\|$MAX$_{\|}\|$ 1	$\mathcal{H}$
a. $H_{0.8}$ $H_{0.8}$ –	-0.8				-21.6
b. $H_{0.8}$ Fus			-1.0		-17
c. $H_{0.8}$ Del		-1.0		-1.0	-21
☞ d. $H_{0.8}$ $H_{0.8}$ Ret				-1	-1

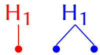
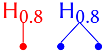
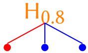
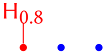
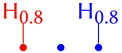
Multiply linked H1: Retraction is always cheaper (Phrase)

3	$H_{0.6}$ $H_{0.6}$	OCP• 27	MAX$_H$ 20	‖UNIF‖ 17	‖MAX$_{\mid}$‖ 1	$\mathcal{H}$
a.	$H_{0.4}$ $H_{0.4}$ –	-0.4				-10.8
b.	$H_{0.4}$ Fus			-1.0		-17
c.	$H_{0.4}$ Del		-0.6		-1.0	-13
☞ d.	$H_{0.4}$ $H_{0.4}$ Ret				-1	-1

Multiply linked H2: Retraction always impossible

(26) $\|\textsc{Max}_{\#|}\|$: Assign -1 violation for every input association between elements A and B that is not realized in the output if there is no other input association of A to X and X precedes B.

1

	$\|\textsc{Max}_{\#\|}\|$ ∞	OCP• 27	$\textsc{Max}_{H}$ 20	$\|\textsc{Unif}\|$ 17	$\|\textsc{Max}_{\|}\|$ 1	$\mathcal{H}$
a.  –		-0.8				-21.6
☞ b.  Fus				-1.0		-17
c.  Del			-1.0		-1.0	-21
d.  Ret	-1				-1	$-\infty$

Lexical exception: No trigger

- object prefix /rí̠/ contains a weakly active $H_{0.6}$ in its underlying representation
  ➙ not strong enough to trigger an OCP repair at 1

1 No fusion triggered by /rí̠/

$H_{0.6}$ $H_{1.0}$		OCP• 27	MAX$_H$ 20	‖UNIF‖ 17	*H 10	SPEC 8	$\mathcal{H}$
☞ a. $H_{0.4}$ $H_{0.8}$	–	-0.4			-1.2	-0.8	-29.2
b. $H_{0.8}$	Del		-1.0		-0.8	-1.2	-37.6
c. –$H_{0.6}$–	Fus			-1.0	-1.2	-0.8	-35.4

Lexical exception: But a regular undergoer

② Deletion predicted for /rí/

$H_{0.8}$ $H_{0.4}$ $H_{0.8}$			OCP• 27	MAX$_H$ 20	‖UNIF‖ 17	*H 10	SPEC 8	$\mathcal{H}$
a.	$H_{0.6}$ $H_{0.2}$ $H_{0.6}$	–	-0.8			-1.4	-1.6	-48.4
☞ b.	$H_{0.6}$ • $H_{0.6}$	Del		-0.4		-1.2	-1.8	-34.4
c.	$H_{0.6}$ $H_{0.2}$ •	Del	-0.6	-0.2		-0.8	-2.2	-45.8
d.	–$H_{0.7}$–	Fus			-2.0	-2.1	-0.9	-62.2

Shona and HLT: Summary

- **consistent decay** of activity across layers predicts different behaviours from a **single grammar**
  - two different spreading domains
  - three different OCP responses
- a **lexical exception** cross-cuts this pattern: predicted from a different underlying activity

HLT and typology of OCP repairs

More HLT accounts of OCP asymmetries

OCP resolutions and OT

H H	OCP_H	$\text{OCP}_\bullet$	Max_H	DepL	‖Unif‖	*!H	Spec
a. H H –	-1	-1					
b. H Del			-1				-1
c. H Fus					-1		
d. H L H DS		-1		-1		-1	
e. H L OW			-1		-1		

Kishambaa: OCP repairs (Odden, 1982)

(27) Two optimal strategies

Input	Optimal at 1 Fus	Optimal at 2 & 3 DS
H H	H	H L H
	–H–	H (L) H

HLT account in a nutshell: Gradual weakening by 0.2

- for a strong H, the downstep repair is blocked since the *!H violations are quite severe
- for weaker H's, the *!H violations are tolerated and L-insertion is a cheaper repair than fusion

Kishambaa: HLT tableaux

1

$H_{1.0}$ $H_{1.0}$			OCP$_H$ 14	*!H 10	‖UNIF‖ 8	‖DEP$_L$‖ 1	$\mathcal{H}$
	a. $H_{0.8}$ $H_{0.8}$	–	-0.8				-11.2
	b. $H_{0.8}$ (L_1) $H_{0.8}$	DS		-0.8		-1.0	-9.8
☞	c. –$H_{0.8}$–	Fus			-1.0		-8

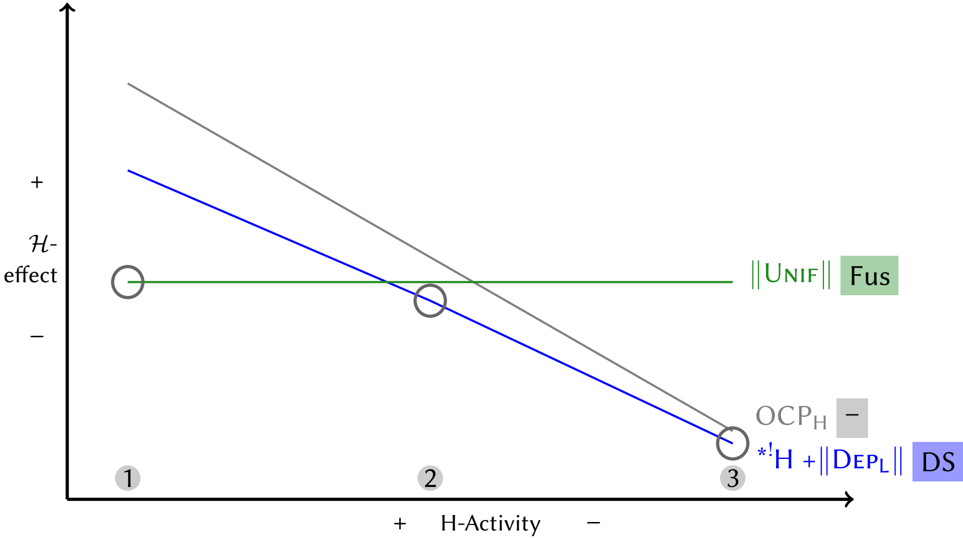
2

$H_{0.8}$ $H_{0.8}$			OCP$_H$ 14	*!H 10	‖UNIF‖ 8	‖DEP$_L$‖ 1	$\mathcal{H}$
	a. $H_{0.6}$ $H_{0.6}$	–	-0.6				-8.4
☞	b. $H_{0.6}$ (L_1) $H_{0.6}$	DS		-0.6		-1.0	-7.7
	c. –$H_{0.6}$–	Fus			-1.0		-8

3

$H_{0.6}$ $H_{0.6}$			OCP$_H$ 14	*!H 10	‖UNIF‖ 8	‖DEP$_L$‖ 0	$\mathcal{H}$
	a. $H_{0.4}$ $H_{0.4}$	–	-0.4				-5.6
☞	b. $H_{0.4}$ (L_1) $H_{0.4}$	DS		-0.4		-1.0	-5.4
	c. –$H_{0.4}$–	Fus			-1.0		-8

Kishambaa: HLT thresholds



Dagaare: OCP repairs (Anttila and Bodomo, 2022)

(28) Three optimal strategies

Input	Optimal at 1 OW	Optimal at 2 DS	Optimal at 3 –
H H	H L	H L H	H H
	H L	H (L) H	H H

HLT account in a nutshell: Gradual weakening by 0.2

- OCP_H only resolved by OW (=strongest H)
- weaker tones only avoid a violation of (29-b) via DS
- the weakest tones don't need to resolve any of the OCP violations

(29) a. OCP_H: Assign -x violation for every H_x followed by another H.
b. $OCP_\bullet$: Assign -x violation for every H_x associated to a TBU that is adjacent to a TBU also associated to a H.

Dagaare: HLT tableaux

1

$H_{1.0}$ $H_{1.0}$		OCP_H 9	OCP• 9	$\|\text{MAX}_H\|$ 6	$\|\text{DEP}_L\|$ 5	$\mathcal{H}$
a. $H_{0.8}$ $H_{0.8}$	–	-0.8	-0.8			-14.4
b. $H_{0.8}$ (L_1) $H_{0.8}$	DS		-0.8		-1.0	-12.2
☞ c. $H_{0.8}$ L_1 ($H_{0.8}$)	OW			-1.0	-1.0	-11

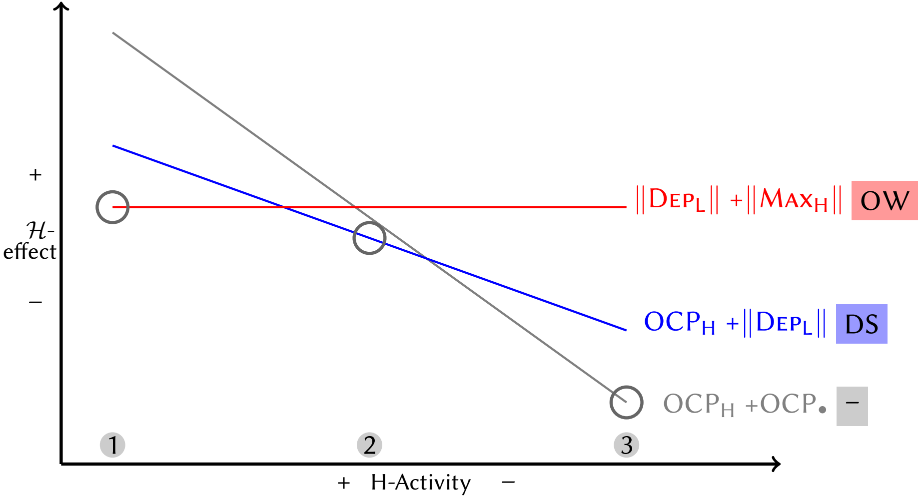
2

$H_{0.8}$ $H_{0.8}$		OCP_H 10	OCP• 9	$\|\text{MAX}_H\|$ 2	$\|\text{DEP}_L\|$ 1	$\mathcal{H}$
a. $H_{0.6}$ $H_{0.6}$	–	-0.6	-0.6			-10.8
☞ b. $H_{0.6}$ (L_1) $H_{0.6}$	DS		-0.6		-1.0	-10.4
c. $H_{0.6}$ L_1 ($H_{0.6}$)	OW			-1.0	-1.0	-11

3

$H_{0.6}$ $H_{0.6}$		OCP_H 10	OCP• 9	$\|\text{MAX}_H\|$ 2	$\|\text{DEP}_L\|$ 1	$\mathcal{H}$
☞ a. $H_{0.4}$ $H_{0.4}$	–	-0.4	-0.4			-7.2
b. $H_{0.4}$ (L_1) $H_{0.4}$	DS		-0.4		-1.0	-8.6
c. $H_{0.4}$ L_1 ($H_{0.4}$)	OW			-1.0	-1.0	-11

Dagaare: HLT thresholds



Bari: OCP repairs (Yokwe, 1986)

(30) Two optimal strategies

Input	Optimal at 1 –	Optimal at 2 & 3 OW
H H	H H	H L
	H H	H L

HLT account in a nutshell: Gradual weakening by 0.2

- $\text{M}\textsc{ax}_{\text{H}}$ is gradient but the relevant ‖OCP•‖ is categorical and has a lower weight
  ➙ Only weakened tones can be deleted to avoid an ‖OCP•‖ violation

Bari: HLT tableaux

1

$H_{1.0}$ $H_{1.0}$		‖UNIF‖ 11	‖OCP•‖ 10	MAX$_H$ 10	‖DEP$_L$‖ 1	$\mathcal{H}$
☞ a. $H_{0.8}$ $H_{0.8}$	–		-1.0			-10
b. –$H_{0.8}$–	Fus	-1.0				-11
c. $H_{0.8}$ L_1	OW			-1.0	-1.0	-11

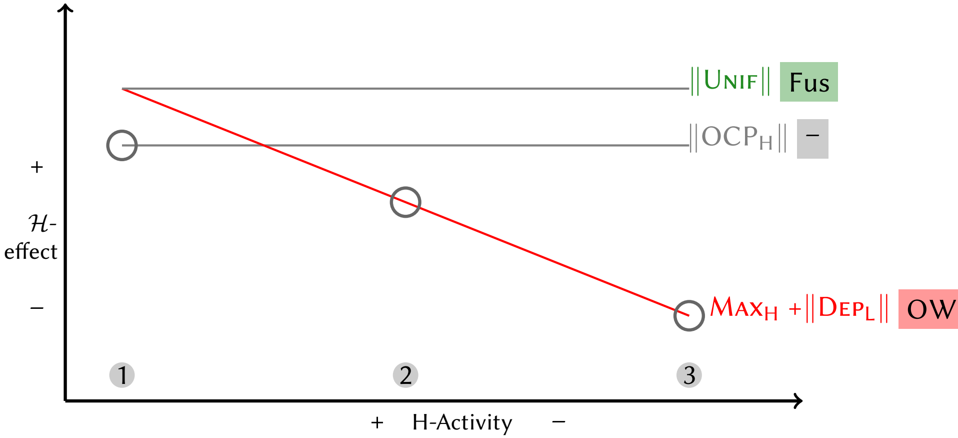
2

$H_{0.8}$ $H_{0.8}$		‖UNIF‖ 11	‖OCP•‖ 10	MAX$_H$ 10	‖DEP$_L$‖ 1	$\mathcal{H}$
a. $H_{0.6}$ $H_{0.6}$	–		-1.0			-10
b. –$H_{0.6}$–	Fus	-1.0				-11
☞ c. $H_{0.6}$ L_1	OW			-0.8	-1.0	-9

3

$H_{0.6}$ $H_{0.6}$		‖UNIF‖ 11	‖OCP•‖ 10	MAX$_H$ 10	‖DEP$_L$‖ 1	$\mathcal{H}$
a. $H_{0.4}$ $H_{0.4}$	–		-1.0			-10
b. –$H_{0.4}$–	Fus	-1.0				-11
☞ c. $H_{0.4}$ L_1	OW			-0.6	-1.0	-7

Bari: HLT thresholds



Tswana: HLT thresholds (Zerbian and Kügler, 2023)

(31) Two optimal strategies

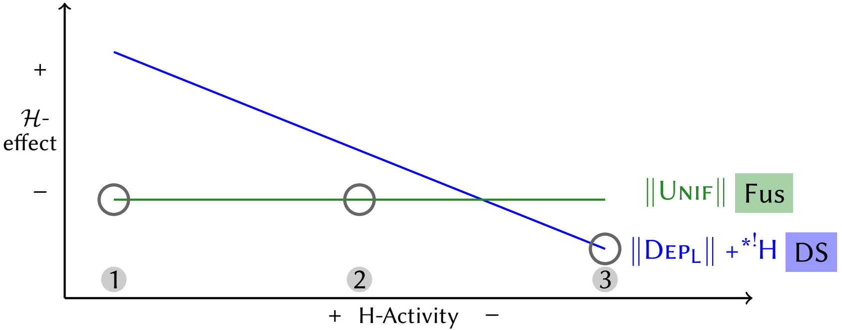
Input	1 + 2 Fus	3 DS
H H	H	H L H

(32) Crucial weights

a. *!H = 10

b. ‖UNIF‖ = 9

c. ‖DEP_L‖ = 2



Tiriki: HLT thresholds (Paster and Kim, 2011)

(33) Two optimal strategies

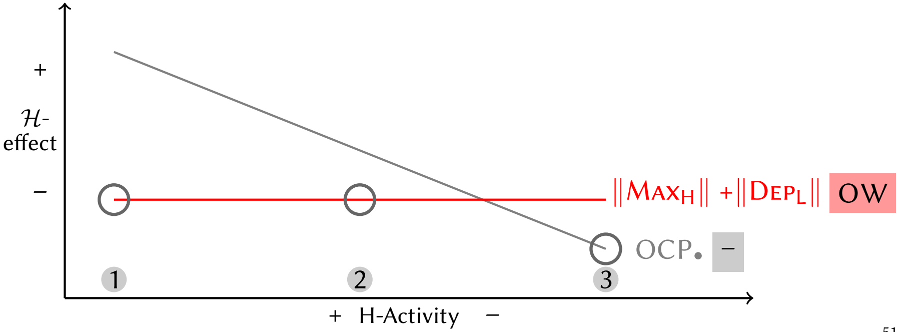
Input	1 + 2 OW	3 –
H H	H L	H H

(34) Crucial weights

a. OCP• = 10

b. $\|$Max$_H\|$ = 4

c. $\|$Dep$_L\|$ = 3



OCP resolutions across layers: Summary of patterns

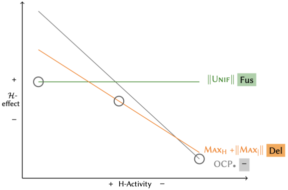
(35)

	Moro (Kordofanian; Sudan)	Shona (East Bantu->Southern Bantu; Mozambique)	Chitonga (East Bantu -> Botatwa; Zambia)	Tiriki (East Bantu->Northeast Savanna Bantu; Kenya)	Dagaare (Gur->Central Gur; Ghana)	Kishambaa (East Bantu->Northeast Savanna Bantu; Tanzania)	Tswana (East Bantu->Southern Bantu; Botswana)	Bari (Nilotic->Eatern Nilotic; South Sudan)
1	Del	Fus	Fus	OW	OW	Fus	Fus	–
2	–	Del	Del	OW	DS	DS	Fus	OW
3	–	–	–	–	DS	DS	DS	OW

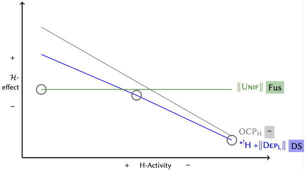
Patterns excluded by HLT

Thresholds in HLT

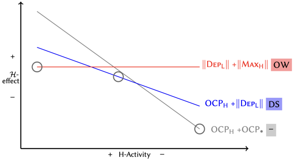
Shona: HLT thresholds



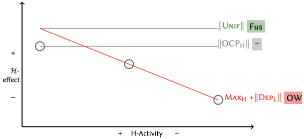
Kishambaa: HLT thresholds



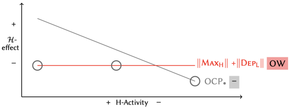
Dagaare: HLT thresholds



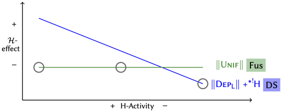
Bari: HLT thresholds



Tiriki: HLT thresholds



Tswana: HLT thresholds



The $\mathcal{H}$-effects of constraints can cross

- if one is categorical and the other gradient
- if two gradient constraints differ in their weighting (='steepness' of their $\mathcal{H}$-effect)

HLT is restricted by monotonicity

Representations become monotonically weaker
+ single constant grammar
= monotonicity of phonological behaviour

(36) Monotonicity of thresholds for phonological behavior in HLT

T_x → Phonological behavior 1

WEAKER: THRESHOLD 1

$T_{x\text{-}y}$ → Phonological behavior 2

WEAKER: THRESHOLD 2

$T_{x\text{-}y\text{-}z}$ → Phonological behavior 3

Monotonicity excludes XYX pattern

(37) Attested OCP repairs across layers

	Moro	Shona	Chitonga	Tiriki	Dagaare	Kishambaa	Tswana	Bari
1	Del	Fus	Fus	OW	OW	Fus	Fus	–
2	–	Del	Del	OW	DS	DS	Fus	OW
3	–	–	–	–	–	DS	DS	OW

(38) Unattested OCP repairs across layers: XYX-patterns

	*Monga	*Shori	*Ciriki	*Tirire	*Dagaba	*Kishana	*Tswari	*Barimba
1	Del	Fus	OW	OW	OW	Fus	Fus	–
2	–	Del	Del	–	DS	DS	Del	OW
3	Del	Fus	OW	OW	OW	Fus	Fus	–

No XYX patterns: Weighting paradox

(39) *Shori in HLT: Impossible

		Necessary weighting relations		
1	Fus	$\textsc{Max}_H$	>	$\lVert\textsc{Unif}\rVert$
2	Del	$\lVert\textsc{Unif}\rVert$	>	$0.8\text{x}\textsc{Max}_H$
3	Fus	$0.6\text{x}\textsc{Max}_H$	>	$\lVert\textsc{Unif}\rVert$

(40) *Tirire in HLT: Impossible

		Necessary weighting relations		
1	OW	0.8xOCP_H	>	$\lVert\textsc{Max}_H\rVert + \lVert\textsc{Dep}_L\rVert$
2	–	$\lVert\textsc{Max}_H\rVert + \lVert\textsc{Dep}_L\rVert$	>	0.6xOCP_H
3	OW	0.4xOCP_H	>	$\lVert\textsc{Max}_H\rVert + \lVert\textsc{Dep}_L\rVert$

HLT vs. stratal OT

(41) *Shori in stratal OT: Easily possible

		Necessary rankings		
1	Fus	OCP, MAXH	≫	UNIF
2	Del	OCP, UNIF	≫	MAXH
3	Fus	OCP, MAXH	≫	UNIF

(42) *Tirire in stratal OT: Easily possible

		Necessary rankings		
1	OW	OCP	≫	MAXH, DEPL
2	–	MAXH, DEPL	≫	OCP
3	OW	OCP	≫	MAXH, DEPL

Conclusion

Conclusion and outlook

- the representative typology of different OCP repairs across layers is predicted under HLT whereas unattested XYX patterns are excluded
- the next questions:
  - Are XYX patterns indeed unattested?
    (e.g. vowel deletion in Seri (Cole, 1986)? – Trommer (to appear) reanalysis as ghost vowels)
  - Is this the best version of HLT?
    (e.g. Zerbian and Kügler (2023) on Tswana: no DS between phonological phrases and four levels)
- will be answered by:
  - extending the OCP typology
  - extending the HLT argument to other tonal behaviour: Spreading, morphological tone realization

References I

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

Anttila, Arto and Adams Bodomo (2022), 'Tone and morphological level ordering in Dagaare', *Phonology* **39**, 443–471.

Bickmore, Lee and Winfred Mkochi (2018), 'OCP effects in Malawian CiTonga tone patterns', *Nordic Journal of African Studies* **27**.

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

Cole, Jennifer (1986), The interaction of phonology and morphology in Seri, *in* 'Proceedings of NELS 17'.

Corina, David P. (1994), The induction of prosodic constraints, *in* S. D.Lima, R.Corrigan and G.Iverson, eds, 'The Reality of Linguistic Rules', John Benjamins, pp. 115–145.

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

Garde, Paul (1965), 'Accentuation et morphologie', *La Linguistique* **1**, 25–39.

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.

References II

Koster, Jan (1986), 'The relation between pro-drop, scrambling, and verb movements', Ms., Rijksuniversiteit Groningen.

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

Kushnir, Yuriy (2019), Prosodic patterns in Lithuanian morphology, PhD thesis, Leipzig University, Leipzig.

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.

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

Myers, Scott (1986), Tone and the Structure of Words in Shona, PhD thesis, University of Massachusetts, Amherst.

Myers, Scott (1997), 'OCP effects in Optimality Theory', *Natural Language and Linguistic Theory* **15**(4), 847–892.

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.

Odden, David (1982), 'Tonal phenomena in Kishambaa', *Studies in African Linguistics* **13**, 177–208.

Paster, Mary and Yuni Kim (2011), 'Downstep in Tiriki', *Linguistic Discovery* **9**, 71–104.

References III

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.

Rizzi, Luigi (1986), 'Null objects in Italian and the theory of pro', *Linguistic Inquiry* **17**, 501–57.

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.

Rosen, Eric (2019*a*), Evidence for gradient input features from Sino-Japanese compound accent, *in* 'Proceedings of AMP 2018', LSA.

Rosen, Eric (2019*b*), 'Learning complex inflectional paradigms through blended gradient inputs', Proceedings of the 2019 meeting of the Society for Computation in Linguistics, January 2019, NYU.

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

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.

Trommer, Jochen (to appear), Strata, *in* 'Cambridge Handbook of Phonology', Cambridge University Press.

References IV

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.

Yokwe, Eluzai M. (1986), The tonal grammar of Bari, PhD thesis, University of Illinois at Urbana-Champaign.

Zerbian, Sabine (2015), 'Adjacent high tones at word- and phrase level in Tswana', talk at the workshop 'Morphosyntactic Triggers of Tone: New Data and Theories', Leipzig University.

Zerbian, Sabine and Frank Kügler (2023), 'Sequences of high tones across word boundaries in Tswana', *Journal of the International Phonetic Association* **53**.

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.

References V

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 (2018*d*), Gradient symbolic representations in the output: A case study from Moses Columbian Salishan stress, *in* S.Hucklebridge and M.Nelson, eds, 'Proceedings of NELS 48', pp. 275–284.

Zimmermann, Eva (2019), Gradient symbolic representations and the typology of ghost segments, *in* K.Hout, A.Mai, A.McCollum, S.Rose and M.Zaslansky, eds, 'Proceedings of AMP 2018', LSA, https://doi.org/10.3765/amp.

Zimmermann, Eva (2021), 'Faded copies: Reduplication as distribution of activity', *Glossa* **6**, 58.

Zimmermann, Eva (to appear), 'Cooperation at the morpho-phonology interface: An argument for phonological locality', *Proceedings of WCCFL 2023* .

Appendix

24. Mai 2024

Bari I

- Yokwe (1986)
- Stem Level:
  - No H-dissimilation
  - No Local H-Spreading
  - Toneless σ's trigger unbounded spreading
  - morphemes:
    - All segmental verb derivation
    - Some tonal verb derivation
    - Singulative -tat
    - Plural -a
- Word Level:
  - H-dissimilation
  - No Local H-Spreading
  - Toneless σ's trigger polar epenthesis
  - morphemes:
    - Associative prefixes

Bari II

- Some tonal verb derivation
- Plural -jin
- Instrumental plural -iʔ

- Phrase Level: Toneless σ's get default L-tone
  - H-dissimilation
  - Local H-Spreading
  - (No toneless σ's?)
  - morphemes: Verbal proclitics (e.g., Future)
- dissimilation is indeed insertion of L since relevant difference between L-toned and toneless TBU's

(43) No OCP-repair at 1
/bójí-tát/
bójí-tát
'trap net'

Bari III

(44) L-Overwriting at 2
/dép-es-íʔ/
hold-INSTR-PL
dép-és-ìʔ/

(45) L-Overwriting at 3
/dók kópò/
dók kòpò
'fetched cup'

Tiriki I

- Paster and Kim (2011)
-

Chitonga I

- Bickmore and Mkochi (2018)
- word level: H-doubling (not across words)
- 3 H-doubling even creates OCP violations

Chitonga: OCP repairs

(46) Three optimal strategies

Input	Optimal at 1 Fus	Optimal at 2 Del	Optimal at 3 –
H H	H	H	H H
	• H	H •	H H

HLT account in a nutshell: Gradual weakening by 0.2

- fusion is the default repair for an OCP violation; in contrast to Shona, the fused H only surfaces on one TBU (due to ‖*LongH‖ +AlignR)
- if the tones get weaker, deletion becomes possible (and ‖Del2!‖ favours deletion of the second H)
- the weakest tone doesn't trigger any OCP repair

Chitonga: HLT tableaux

1

	$H_{1.0}$ $H_{1.0}$		OCP• 10	Max_H 8	$\lVert$Unif$\rVert$ 5	$\lVert \text{Max}_l \rVert$ 1	$\mathcal{H}$
a.	$H_{0.8}$ $H_{0.8}$	–	-0.8				-8
☞ b.	• $H_{0.8}$	Fus			-1.0	-1.0	-6
c.	$H_{0.8}$ •	Del		-1.0		-1.0	-9

2

	$H_{0.8}$ $H_{0.8}$		OCP• 10	Max_H 8	$\lVert$Unif$\rVert$ 5	$\lVert \text{Max}_l \rVert$ 1	$\mathcal{H}$
a.	$H_{0.6}$ $H_{0.6}$	–	-0.8				-8
☞ b.	• $H_{0.6}$	Fus			-1.0	-1.0	-6
c.	$H_{0.6}$ •	Del		-1.0		-1.0	-9

2

	$H_{0.6}$ $H_{0.6}$		OCP• 10	Max_H 8	$\lVert$Unif$\rVert$ 5	$\lVert \text{Max}_l \rVert$ 1	$\mathcal{H}$
a.	$H_{0.4}$ $H_{0.4}$	–	-0.8				-8
☞ b.	• $H_{0.4}$	Fus			-1.0	-1.0	-6
c.	$H_{0.4}$ •	Del		-1.0		-1.0	-9

Tswana I

- Zerbian and Kügler (2023):
  - ‘sequences of singly-linked adjacent high tones are allowed but sequences of multiply-linked adjacent high tones are not allowed (Mmusi 1992)’
  - Mmusi suggests that separate lexical H’s are fused
  - OCP also blocks HSpread
    (it’s not about underlyingly multiply linked tones!?)
  - OCP-violations above the word are repaired by DS but only within phonological phrases (their study!)
    (more data from Southern Sotho)
- Zerbian (2015)
  - H on subject/object prefix and stem are realized
  - ‘across certain word-internal junctures’, however, deletion takes place; resp. ‘left-branch delinking’
- ZerbianKuegler:2015

Tswana II

- the OCP causes 'either the blocking of high tone spread [...], the fusion of two singly-linked tones [...] or the delinking of a high tone'

(47) Fusion within words

a. ó-réká nama
sc1-buy meat
'She is buying meat' (Zerbian and Kügler, 2023, 364)

b. go-é-rékisa
sc15-oc9-sell
'to sell it' (Zerbian, 2015)

(48) Deletion within words

a. ó-ká-reká
sc1-ASP-buy
'S/he is able to buy' (Zerbian, 2015)

(49) H-Spread blocked

Tswana III

a. ó-a-réká
SC1-ASP-buy
'She is buying' (Zerbian and Kügler, 2023, 364)

(50) Downstep across words (within phrase)

a. Ba-rwá ꜜbá-thúsá ba-lemi
NP2-son SC2-help NP2-farmer
'The sons help the farmers' (Zerbian and Kügler, 2023, 368)

b. Ba-ńná ꜜbá-a-húma
NP2-man SC2-ASP-become.rich
'The men are becoming rich' (Zerbian and Kügler, 2023, 368)

Kishambaa

- Atlantic-Congo ➝ East Bantu ➝ Northeast Savanna Bantu
- spoken in Tanzania
- sources for data and generalizations: Myers (1997)

Kishambaa: OCP data

(51) Fusion at 1 (and spread at 3), p.883

ni[kááng-é] nyáma	cf. ku-[kááng-a]
1stsg-fry-perfect meat	*infinitive-fry-term*
I fried meat	to fry

[kááng-á]
fry-term
Fry!

(52) OCP tolerated at 3, p.880

a.	nwáná 'child'	dú 'only'	nwánáꜝdú 'only a child'
b.	kúi 'dog'	ní 'cop'	níꜝkúi 'it is a dog'

(53) OCP tolerated at (2 /) 3, fed by spreading, p.883

ní-kí-[ꜝchí-kómá]
1stsg.-progressive-it-kill
I was killing it (Cl. 7)

Kishambaa: Spreading data

(54) Spreading at (2/) 3, p.880

a. ku-ví̲-ghóshó-a — cf. ku-ghosho-a
to do them (Cl. 8) — *to do*

b. ni-té̲-ghóshó-a
I have done (verb-focus)

(55) Spreading at 3, p.880

nyumbá̲ zá-wá-ghánga — cf. za-wa-ghanga
house of the doctors — *of the doctors*

(56) Spreading at 3, creating OCP-violation (=DS), p.881

ni-on-í̲yé̲ má-!kú̲i — cf. ni-on-í̲yé̲, — ma-kú̲i
I saw dogs — *1stsg-see-perfect* — *6-dog*
I saw

ni-on-í̲yé̲ nyú!mbá̲ — cf. ni-on-í̲yé̲, — nyumbá̲
I saw the house — *1stsg-see-perfect* — house
I saw

Kishambaa: Object prefixes

(57) No fusion for Obj in penult/stressed position at [2], p.883

ku-[chí̱-ꜜjá̱]
infinitive-it(7)-eat
to eat it (Class 7)

ni-té̱-[ꜜí̱-ꜜnywá̱]
1stsg-past-it(9)-drink
I drank it (Class 9)

(58) Fusion for Obj in unstressed position at [2], p.883

ku-[wá̱-kó̱má]	cf. ku-[kó̱má]
infinitive-them-kill	*infinitive-kill*
to kill them	to kill

ní̱-kí-[ꜜchí̱-kó̱má]
1stsg.-progressive-it-kill
I was killing it (Cl. 7)

Kishambaa: Subjunctive

(59) Lowering in the subjunctive, 2, p.884

né ní-[wa-kaang-iy-e]	cf. ku-[kááng-a]
I will fry for them (soon)	to fry
ní-[wa-kom-e]	cf. ku-[kóm-á]
I should kill them	to kill

Constraints

(60) $*^{!}$H: Assign -x violation for every H_x preceded by a floating L.

(61) $\|$UNIF$\|$: Assign -1 violation for every pair of input tones corresponding to the same output tone.

GSR: Broader Context

- that linguistic elements are not categorical but have strength differences is **not a new** idea
  (e.g. Rizzi (1986) and Koster (1986) for functional categories in syntax, Garde (1965): some lexical accent system are based on scalar grades of accent strength,...)
- other work on non-categorical elements in **neural networks**
  (e.g. Corina (1994) on induction of prosodic categories in neural networks)
- can also predict **phonetic gradience**
  (e.g. subphonemic gradience in word-final devoicing, nasal place assimilation, flapping (e.g. Braver, 2013), vowel harmony is gradient (McCollum, 2018),...)
- **different from a binary** distinction into strong/weak
  (Inkelas, 2015; Vaxman, 2016*a*,*b*; Sande, 2017)

→ here: predictions of **gradient (=numerical) phonological strength in an OT-system** as explanation for 'exceptional' behaviour

General 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, 2018*d*)
   - 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)
   - (interacting) ghost segments in Welsh (Zimmermann, 2018*c*)
   - ...