RUTGERS

UNIVERSITY | NEW BRUNSWICK

Computation of unbounded processes with a suprabinary scale requires non-determinism

Nate Koser and Adam McCollum
Rutgers University
NELS 2021

Introduction

- What is the computational nature of stress assignment?[1]
  - $\sigma\sigma\sigma\sigma\sigma\sigma \mapsto \acute{\sigma}\sigma\grave{\sigma}\sigma\grave{\sigma}\sigma$
- What kind of *functions* are these?
- Intuition that some are more complex than others:
  - initial stress: $\acute{\sigma}, \acute{\sigma}\sigma, \acute{\sigma}\sigma\sigma, \acute{\sigma}\sigma\sigma\sigma$... vs.
    default-to-opposite: LLLH́L, HHHHH́, ĹLLLL

[1] I focus here on patterns where stress is predictable, rather than lexically-conditioned

Introduction

- Formal language theory (FLT) gives rigorous measure of complexity
- Nested hierarchy of function classes based on their expressivity
  - initial stress: simple vs.
    default-to-opposite: more complex
- Gives formal definition of *locality*
- Leads to testable hypotheses about stress typology and its learnability[2]

[2]For some previous work, see Heinz (2009); Rogers et al. (2013); Baek (2018); Hao & Anderson (2019); Koser & Jardine (2020)

Takeaway

- Focus today on non-local (unbounded) patterns
- Demonstrate that complexity affected not only by sensitivity to the input/output or locality, but the **number of contrasts**
- Complex patterns break down into steps that are maximally *piecewise*[3]

[3]languages: Rogers et al. (2010); functions: Burness & McMullin (2020)

Results: empirical

- binary weight scale (DTS): $\overrightarrow{\text{LLLLHHL}}$

  suprabinary scale: $\overrightarrow{\text{LMLMLHLL}}$

- With a binary scale, presence of H terminates search for stress

  - can be computed deterministically

- With a suprabinary scale, must posit syllables of intermediate weight as possible stress targets

  - *cannot* be computed deterministically, more complex

Results: empirical

- Phonological function with *unary* input (QI stress; σ^n) *cannot* describe a long-distance pattern
- binary weight scale (DTS): $\overrightarrow{\text{LLLLH}}\text{HL}$

  unary input as in QI stress: $\overrightarrow{\sigma\sigma\sigma\sigma\sigma\sigma\sigma}$
- Offers explanation as to why unbounded patterns are only ever QS
- Number of scalar levels and size of input inventory significantly alter formal complexity of phonological generalizations

Results: theoretical

- Can view patterns as *decomposition* into separate functions
- Important that steps of decomposition are principled – many extensionally equivalent I-O maps, what do they say about *phonology*
- Result: *piecewise* restriction on individual functions i.e. individual phonological generalizations

Results: theoretical

- Piecewise function – consistent effect of inputs on further inputs
  - stress leftmost H: LH́LLHH, LLLH́HH, LLH́LLLH
- **No** "meta" properties e.g. parity counting – increase in formal complexity
  - stress alternating H: H́LLHH́LHH́, H́LLLLHH́, LLH́HH́LL
- Being *phonological* means referring only to symbols in the word

Plan

- Background
- Data and Analyses
- Discussion

Complexity

- FLT complexity classes divide space of possible functions based on expressive power of those functions
- Phonology is *regular*[4] – definable with a finite-state machine
- In fact, most is subregular[5]
- Something intuitively non-phonological about English center embedding, this approach tells us why

[4]Johnson (1972); Kaplan & Kay (1994)

[5]Rogers et al. (2013); Heinz (2018)

Complexity

- Within unbounded typology, some relevant complexity distinctions:
- subsequential functions[6] – any one-way deterministic function
    - stress alternating heavy σ: ĹLLHĹLHĹ, ĹLLLHĹ, LLĹHHĹLL...
- strictly piecewise (SP) functions[7] – consistent effect of previous inputs on further inputs
    - stress H in DTO: LĹLLHH, ĹHHHH, LLĹLLLH

[6]Mohri (1997)

[7]Burness & McMullin (2020)

Decomposition

- Some patterns can't be described in a deterministic one-way manner
- Yana[8] – 'leftmost heavy or left' (LHOL):

(1)
a. p'ú.di.wi 'women'
b. si.búm.k'ai 'sandstone'
c. tsi.ni.já: 'no'

[8]Sapir & Swadesh (1960)

Decomposition

- To locate stress: R-to-L – $\text{LHHH}\overset{?}{\overleftarrow{\text{H}}}\text{LL}\#$

  L-to-R – $\overrightarrow{\#}\overset{?}{\text{L}}\text{HHHLL}$

- View these as *decomposition* into two functions that each express some aspect of the total map

- *Leftmost Heavy* – stress first heavy, left to right

  Or Left – stress initial σ if no heavies, right to left

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV			⋉

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV	CV		⋉

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV	CV	CV	⋉

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV	CV	CV	⋉
Or Left	⋊			CV	⋉

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV	CV	CV	⋉
Or Left	⋊		CV	CV	⋉

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV	CV	CV	⋉
Or Left	⋊	CV́	CV	CV	⋉

Analysis

- Binary LHOL: Yana

/p'udiwi/ $\mapsto$ [p'ú.di.wu]

input	⋊	CV	CV	CV	⋉
Leftmost Heavy	⋊	CV	CV	CV	⋉
Or Left	⋊	CV́	CV	CV	⋉
output	⋊	CV́	CV	CV	⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV			⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV	CV		⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV	CV	CV́:	⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV	CV	CÝ:	⋉
Or Left	⋊			CÝ:	⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV	CV	CV́:	⋉
Or Left	⋊		CV	CV́:	⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV	CV	CV́:	⋉
Or Left	⋊	CV	CV	CV́:	⋉

Analysis

- Binary LHOL: Yana

/tsinija/ $\mapsto$ [tsi.ni.já:]

input	⋊	CV	CV	CV:	⋉
Leftmost Heavy	⋊	CV	CV	CV́:	⋉
Or Left	⋊	CV	CV	CV́:	⋉
output	⋊	CV	CV	CV́:	⋉

Analysis

- For LHOL[9], functions are non-interacting – order doesn't matter
- Non-interacting decompositions less expressive than interacting decompositions – *weakly deterministic*[10]
- *Any* ternary or higher scale in unbounded pattern *requires* interaction

[9] And DTO/DTS in general

[10] Meinhardt et al. (submitted); Heinz & Lai (2013)

Analysis

- Suprabinary LHOL: Mauritanian Pulaar[11]
- quaternary weight distinction[12]: CV:C » CV: » CVC » CV

(2)

a.	á.du.na	'world'	c.	hál.ku.de	'to make perish'
b.	hal.ká:.de	'to perish'	d.	ja:.tá:r.na:.jo	'person from Jaataar'

[11] Niang (1997)

[12] See Gordon (2006, p.126) for other patterns of this type

Analysis

- Suprabinary LHOL: Mauritanian Pulaar – analysis above fails

/halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost heavy	⋊	CV́C			⋉

Analysis

- Suprabinary LHOL: Mauritanian Pulaar

/halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost Heavy	⋊	CV́C	CV:		⋉

Analysis

- Suprabinary LHOL: Mauritanian Pulaar

/halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost Heavy	⋊	CV́C	CV:	CV	⋉

Analysis

- Suprabinary LHOL: Mauritanian Pulaar

/halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost Heavy	⋊	CV́C	CV:	CV	⋉
Or Left	⋊			CV	⋉

Analysis

- Suprabinary LHOL: Mauritanian Pulaar

/halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost Heavy	⋊	CV́C	CV:	CV	⋉
Or Left	⋊		CV:	CV	⋉

Analysis

- Suprabinary LHOL: Mauritanian Pulaar

/halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost Heavy	⋊	CV́C	CV:	CV	⋉
Or Left	⋊	CV́C	CV:	CV	⋉

Analysis

- Suprabinary LHOL: Mauritanian Pulaar

  /halka:de/ $\mapsto$ [hal.ká:.de]

input	⋊	CVC	CV:	CV	⋉
Leftmost Heavy	⋊	CV́C	CV:	CV	⋉
Or Left	⋊	CV́C	CV:	CV	⋉
output	⋊	CV́C	CV:	CV	⋉

- */hal.ka:de/ $\mapsto$ [hál.ka:.de]

What happened?

- Stress leftmost heavi*est*, or left
- With binary scale, first H in word is only potential target
- With suprabinary scale, *every* heaviest syllable seen is potential target

What to do?

- First function marks potential targets, second locates correct one
- *Heaviest* – stress heaviest σ thus far seen, right to left

  Destress – keep first stress, delete others, left to right
- *Heaviest* creates intermediate form where leftmost stress is the one to keep, *Destress* uses this information, enforces culminativity
- Explicit interaction necessitated by suprabinary weight scale

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊				CV́	⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊			CÝ:	CÝ	⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊		CV́:C	CV́:	CV́	⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊	CV:	CÝ:C	CÝ:	CÝ	⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊	CV:	CV́:C	CV́:	CV́	⋉
Destress	⋊	CV:				⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ ↦ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊	CV:	CÝ:C	CÝ:	CÝ	⋉
Destress	⋊	CV:	CÝ:C			⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊	CV:	CV́:C	CV́:	CV́	⋉
Destress	⋊	CV:	CV́:C	CV:		⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ $\mapsto$ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊	CV:	CV́:C	CV́:	CV́	⋉
Destress	⋊	CV:	CV́:C	CV:	CV	⋉

Analysis: Pulaar

- /ja:.ta:r.na:.jo/ ↦ [ja:.tá:r.na:.jo]

input	⋊	CV:	CV:C	CV:	CV	⋉
Heaviest	⋊	CV:	CV́:C	CV́:	CV́	⋉
Destress	⋊	CV:	CV́:C	CV:	CV	⋉
output	⋊	CV:	CV́:C	CV:	CV	⋉

Reflections

- Must interact – order is crucial: *Destress* first leads to outputs like *[ja:.tá:r.ná:.jó]
- Total map is non-deterministic, at *regular* bound on phonology

Reflections

- BUT each individual function expresses a simpler, SP generalization:
  - *Heaviest*: CV́:C...{CV́:C, CV:, CVC, CV}, etc. for other σ
  - *Destress*: $\acute{\sigma}...\sigma$
- Each input has *consistent*, long-distance effect on further inputs
- Resembles other well-known processes, e.g. harmony:
  - Navajo: ʃ...ʃ, *ʃ...s
- If phonology is composition of individual generalizations, suprabinary patterns are not unusual in terms of complexity

Non-phonological?

- Something that's not piecewise?
- "3σ suprabinary LHOL" – Pulaar but only stress in 3σ windows:

(3) a. CV́. CV. CV. CV:. CV:C. CV.

b. CV. CVC. CV. CV. CV:. CV. CV́:

- Not piecewise – *inconsistent* effect of inputs on further inputs

Non-phonological?

- Something that's not piecewise?
- "3σ suprabinary LHOL" – Pulaar but only stress in 3σ windows:

(4) a. CV́. CV. CV. CV:. CV:C. CV.

b. CV. CVC. CV. CV. CV:. CV. CV́:

- Not piecewise – *inconsistent* effect of inputs on further inputs in "3σ *Heaviest*"
  - both CVC...CV: and CVC...CV́:

Non-phonological?

- "3σ *Heaviest*" not piecewise, *is* subsequential
- Suggests that individual pieces of a generalization subject to piecewiseness
- Piecewiseness hinges on symbols in the string, subsequential functions can introduce parity and counting
- Being *phonological* is about the symbols in the string, not meta-properties like parity

Recap

- Significant differences in complexity based on scalar levels
    - Binary scale, unbounded pattern can be deterministic
    - Suprabinary scale necessarily non-deterministic
- Each individual step of the total map of attested patterns still adheres to piecewiseness
- Two more related points...

More levels

- More scalar levels beyond i.e. 3 vs 4 vs 12 does not affect complexity
- Nanti[13]: twelve-step scale combines weight and sonority – Ca:N as best target and Ci as worst target, stress rightmost

[13]Crowhurst & Michael (2005)

More levels

- Despite explosion of scale, location of main stress proceeds as in Pulaar:

  iŋ.kan.tán.ta.ksem.pa.ra 'he will say that for a reason'

  (5) *Heaviest:* CíN CáN CáN Ca CeN Ca Ca
  Destress: CiN CaN CáN Ca CeN Ca Ca

More Levels

- Still piecewise, e.g. observe a Ca, further better targets stressed, worse targets unstressed
- Number of levels beyond suprabinary doesn't further increase complexity
- More labels for positions in word, behavior of function with regards to the word is identical

Unary scale

- Significant differences in computation in binary vs. suprabinary scales
- What about binary vs. unary scale?
    - quantity insensitive (QI) stress inputs: σ^n, unary
- With unary input, cannot define unbounded pattern

Unary scale

(6)

inputs = L,H			inputs = σ		
a. LHLLLL	$\mapsto$	LH́LLLL	a. $\sigma\sigma\sigma$	$\mapsto$	$\acute{\sigma}\sigma\sigma$
b. LLLLHL	$\mapsto$	LLLLH́L	b. $\sigma\sigma\sigma\sigma$	$\mapsto$	$\acute{\sigma}\sigma\sigma\sigma$
c. LLLLLL	$\mapsto$	ĹLLLLL	c. $\sigma\sigma\sigma\sigma\sigma$	$\mapsto$	$\acute{\sigma}\sigma\sigma\sigma\sigma$

- Presence of at least two input symbols allows encoding of a long-distance generalization
- With only one, there's no differences to track, can't encode long-distance generalization[14]

[14]Possible counterexample from Creek stress in all L forms, but this is local assuming foot structure

Unary scale

(6)

inputs = L,H			inputs = σ		
a. LHLLLL	$\mapsto$	LH́LLLL	a. $\sigma\sigma\sigma$	$\mapsto$	$\acute{\sigma}\sigma\sigma$
b. LLLLHL	$\mapsto$	LLLLH́L	b. $\sigma\sigma\sigma\sigma$	$\mapsto$	$\acute{\sigma}\sigma\sigma\sigma$
c. LLLLLL	$\mapsto$	ĹLLLLL	c. $\sigma\sigma\sigma\sigma\sigma$	$\mapsto$	$\acute{\sigma}\sigma\sigma\sigma\sigma$

- Limitation to unary input alphabet restricts the range of possible phonological functions
- Offers explanation as to why unbounded patterns only appear in quantity sensitive (QS) systems

Discussion

- Number and type of contrasts a language makes is a crucial factor in the expressivity of possible patterns
- Suprabinary unbounded patterns require non-determinsitc map
- Not just syllable inventory – the *interaction* of processes *with* the inventory:
    - Language with high, mid, and low vowels, but sonority is irrelevant

Discussion

- Important that steps of decomposition are principled – many extensionally equivalent decompositions, what do they say about *phonology*
- Posit piecewise boundary on individual steps of a decomposition – no reference to meta-properties like parity
- Tracks with previous work on stress as formal languages i.e. stringsets[15] if function composition is analogous to set intersection
- Tells us how the functions can be learned from positive data[16]

[15] Rogers et al. (2013)

[16] Heinz & Rogers (2010)

Thanks!

References

Baek, H. (2018). Computatonal representation of unbounded stress. *Proceedings of CLS 53*, (pp. 13–24).

Burness, P., & McMullin, K. (2020). Modeling non-local maps as Strictly Piecewise functions. *Proceedings of the Society for Computation in Linguistics*, *3*.

Crowhurst, M., & Michael, L. (2005). Footing and prominence-driven stress in Nanti (Kampa). *Language*, *81*(1), 47–95.

Gordon, M. (2006). *Syllable Weight: Phonetics, Phonology, Typology*. New York, NY: Routledge Press.

Hao, S., & Anderson, S. (2019). Unbounded stress in subregular phonology. *Proceedings of SIGMORPHON 16*.

Heinz, J. (2009). On the role of locality in learning stress patterns. *Phonology*, *26*(2), 303–351.

Heinz, J. (2018). The computational nature of phonological generalizations. In L. M. H. . F. Plank (Ed.) *Phonological typology*. Berlin & Boston: De Gruyter Mouton.

Heinz, J., & Lai, R. (2013). Vowel harmony and subsequentiality. In A. Kornai, & M. Kuhlmann (Eds.) *Proceedings of the 13th Meeting on Mathematics of Language*. Sofia, Bulgaria.

Heinz, J., & Rogers, J. (2010). Estimating strictly piecewise distributions. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Uppsala, Sweden: Association for Computational Linguistics.

Johnson, D. (1972). Formal aspects of phonological description.

Kaplan, R., & Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, *20*, 331–378.

Koser, N., & Jardine, A. (2020). The computational nature of stress assignment. *In Proceedings of AMP 2019*.

Meinhardt, E., Mai, A., Bakovic, E., & McCollum, A. (submitted). On the proper treatment of weak determinism: Subsequentiality and simultaneous application in phonological maps. *https://ling.auf.net/lingbuzz/005565*.

Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, *23*, 269–311.

Niang, M. (1997). *Constraints on Pulaar Phonology*. New York: University Press of America.

Rogers, J., Heinz, J., Bailey, G., Edlefsen, M., Visscher, M., Wellcome, D., & Wibel, S. (2010). On languages piecewise testable in the strict sense. In C. Ebert, G. Jäger, & J. Michaelis (Eds.) *The Mathematics of Language*, (pp. 255–265). Berlin, Heidelberg: Springer Berlin Heidelberg.

Rogers, J., Heinz, J., Fero, M., Hurst, J., Lambert, D., & Wibel, S. (2013). Cognitive and sub-regular complexity. In G. Morrill, & M.-J. Nederhof (Eds.) *Formal Grammar*, vol. 8036 of *Lecture Notes in Computer Science*, (pp. 90–108). Springer.

Sapir, E., & Swadesh, M. (1960). *Yana Dictionary*. Berkeley: University of California Press.