

The Stone Unturned

Jeffrey Heinz



Stony Brook University

NELS 57



CUNY

2026 10 16

PROLOGUE

Legend has it that the Greek General Mardonius buried a treasure and Polycrates searched for it without luck. He asked the **Oracle of Delphi** how he could find it and the answer was:

“Πάντα λίθον κίνει”

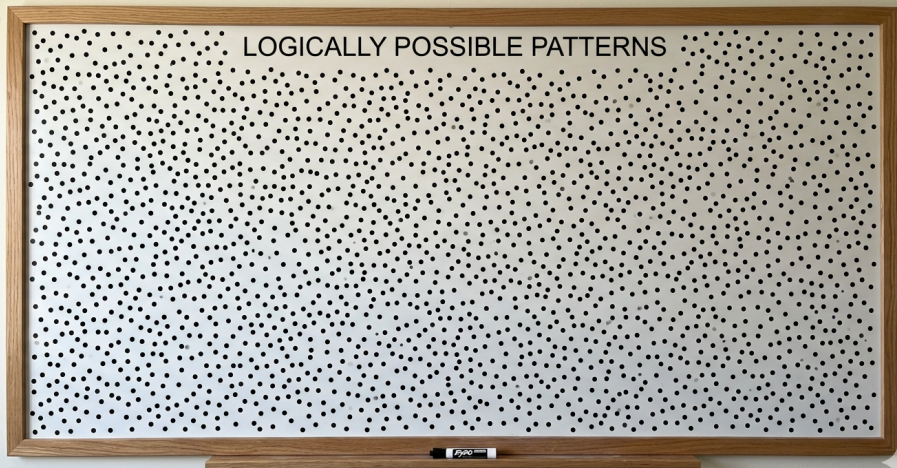
Pánta líthon kínei

Turn every stone.

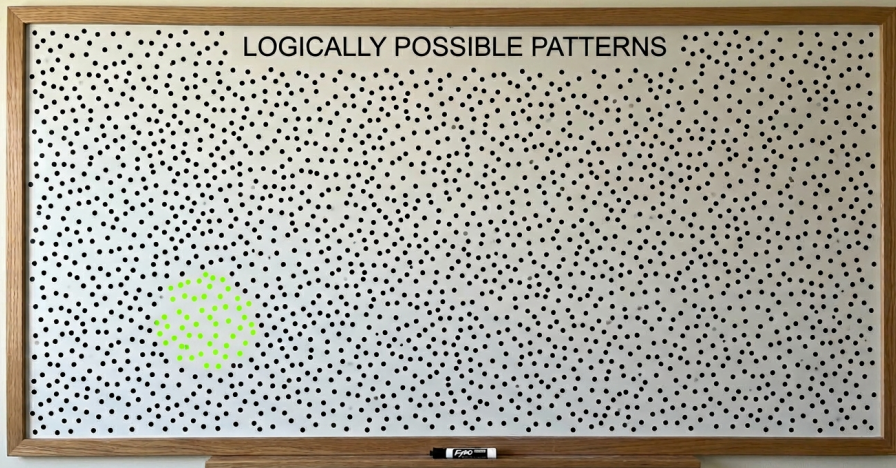
TODAY

- 1 I argue that **symbolic, formal systems** studied in mathematical and computational linguistics remain mostly **unexplored**, and that many rich treasures, which **provide insight into natural languages**, remain waiting to be discovered.
- 2 My argument rests on **grammatical analyses and learning results** based on **fine-grained computational classes**.
- 3 These results are **interpretable, broadly faithful to linguistic theories** (e.g. locality, representations), **learnable from small data**, make cognitive predictions, and more.

CHARACTERIZING NATURAL LANGUAGE



CHARACTERIZING NATURAL LANGUAGE



CHARACTERIZATIONS OF SYNTAX AND PHONOLOGY

Computably Possible

S

Context-Sensitive

P

Regular

Subregular

strings

(Chomsky, 1956; Johnson, 1972; Huybregts, 1984; Shieber, 1985;
Kaplan and Kay, 1994)

CHARACTERIZATIONS OF SYNTAX AND PHONOLOGY

Computationally Possible

S

Context-Sensitive

Regular

P

Subregular

strings

(Heinz, 2018, 2025)

CHARACTERIZATIONS OF SYNTAX AND PHONOLOGY

Computably Possible

Context-Sensitive

S

Regular

P

Subregular

trees strings

(Kobele, 2011; Graf, 2012; Stabler, 2019)

CHARACTERIZATIONS OF SYNTAX AND PHONOLOGY

Computationally Possible

Context-Sensitive

Regular

S

P

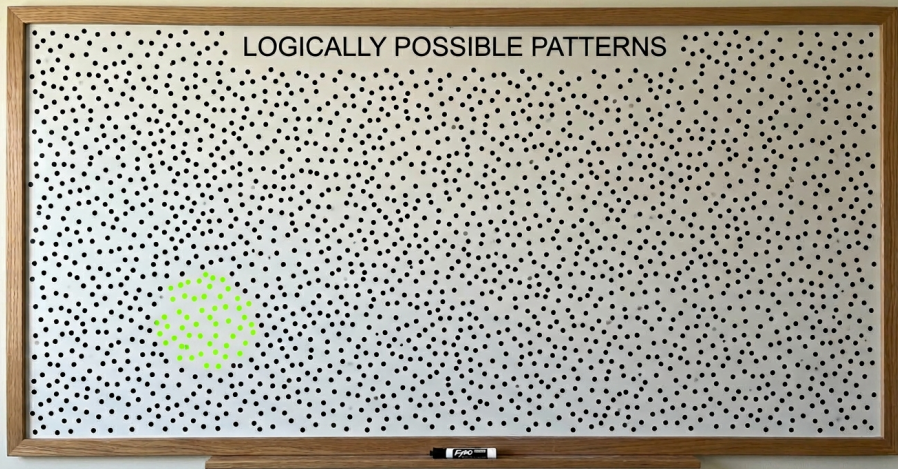
Subregular

trees

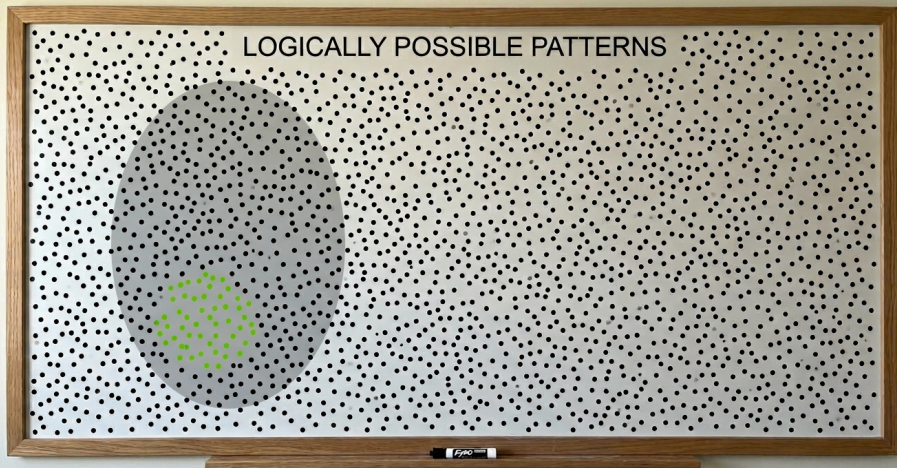
strings

(Graf, 2022)

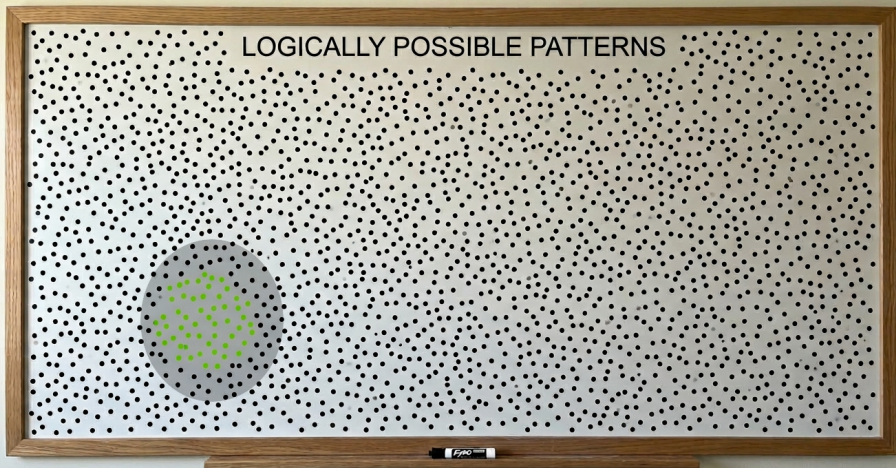
CHARACTERIZING NATURAL LANGUAGE



CHARACTERIZING NATURAL LANGUAGE



CHARACTERIZING NATURAL LANGUAGE



ADVANTAGES OF MORE PRECISE CHARACTERIZATIONS

Mathematical and computational characterizations

- 1 are **grounded** in rich theories of formal grammars, logic, algebra and automata
- 2 are inclusive of **linguistic representations** (strings, trees, features, autosegmental, prosodic, morphological, ...)
- 3 provide methods to **verify** complex analyses of linguistic phenomena
- 4 make clear **typological** predictions
- 5 have led to new **learning algorithms** (which generalize more correctly with less data)
- 6 have **cognitive analogues** and make psycholinguistic predictions
- 7 can be applied to **many subdomains of language** including phonetics, phonology, morphology, syntax, semantics, pragmatics

A SECOND THEME

optimization $\sim$ satisfaction

statistics $\sim$ structure



A SECOND THEME

optimization $\sim$ **satisfaction**

statistics $\sim$ **structure**



A THIRD THEME

Logical languages as a window into natural languages.

Monadic Second Order Logic $\sim$ **Regular**

|
First Order Logic

|
Propositional Logic (Quantifier Free)

|
Conjunctions

WHAT'S NEXT

- 1 **What ‘Regular’ Means**
- 2 Locality/Subregularity in Phonological Processes
- 3 Locality/Subregularity in Phonological Constraints
- 4 Learning Phonological Constraints
- 5 Learning Phonological Processes
- 6 Beyond Phonology
- 7 Conclusion

REGULAR GRAMMARS FOR CONSTRAINTS AND TRANSFORMATIONS

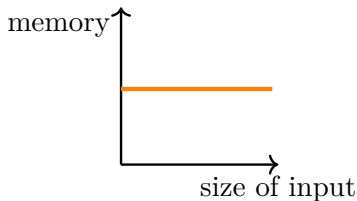
Multiple Equivalent Formalizations of “Regular”

Theory	Result
Logic	Monadic Second Order
Automata	Finite-state
Algebra	Finite Semigroups

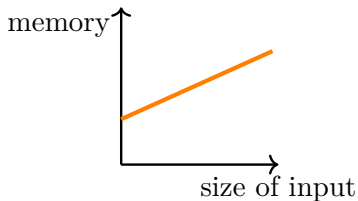
Kleene (1956); Büchi (1960); Pin (2021)

WHAT “REGULAR” MEANS

A grammar is regular provided **the memory required to assemble the computation is bounded by a constant, regardless of the size of the input.**



Regular



Non-Regular

EXAMPLE: VOWEL HARMONY

Progressive

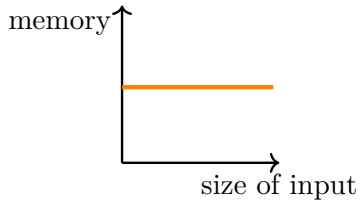
Vowels agree in backness with the first vowel in the underlying representation.

Majority Rules

Vowels agree in backness with the majority of vowels in the underlying representation.

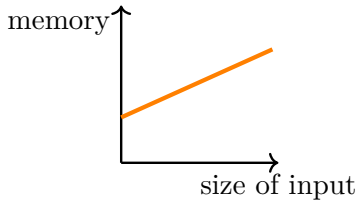
UR	Progressive SR	Majority Rules SR
/no-ke-lu/	nok <u>o</u> lu	nok <u>o</u> lu
/no-ke-li/	nok <u>o</u> lu	n <u>i</u> keli
/pi-du-go/	pid <u>i</u> ge	p <u>u</u> dugo
/pi-dugo-me-mi/	pid <u>i</u> gememi	pid <u>i</u> gememi

PROGRESSIVE AND MAJORITY RULES HARMONY



Regular

Progressive



Non-Regular

Majority Rules

SOME PERSPECTIVE

Typological

Majority Rules is unattested. (Lombardi 1999, Bakovic 2000)

Psychological

Human subjects fail to learn Majority Rules in artificial grammar learning experiments, unlike progressive harmony.
(Finley 2008, 2011)

Computational

Majority Rules is not regular. (Riggle 2004, Heinz and Lai 2013)

OPTIMALITY THEORY

- 1 There exists a CON and ranking over it which generates Majority Rules: $\text{AGREE}(\text{BACK}) \gg \text{IDENTIO}[\text{BACK}]$.
- 2 Some believe changing CON may resolve this, but such an approach fails to recognize the core problem.
- 3 The problem appears to be global optimization itself. There are many reasons to think CON cannot be so changed! (Hao 2019, 2024; Lamont 2021, 2022)

WHAT'S NEXT

- 1 What 'Regular' Means
- 2 **Locality/Subregularity in Phonological Processes**
- 3 Locality/Subregularity in Phonological Constraints
- 4 Learning Phonological Constraints
- 5 Learning Phonological Processes
- 6 Beyond Phonology
- 7 Conclusion

Word-final /e/ raising in Finnish (Uralic, Finland)

/ # r e k e # /

[# r e k **i** #]

‘sledge (nom. sg.)’

Odden (2013)

LOCALITY AS AN ADJACENCY WINDOW

Word-final /e/ raising in Finnish (Uralic, Finland)

/ # r e k e # /

[# r e k i #]

‘sledge (nom. sg.)’

- The max window size is called the ‘k’ parameter. Here $k=2$.
- /e/ becomes [i] if its neighbor to the right is the boundary.
- Let’s call such processes **Strictly Local**.

Odden (2013)

LOCALITY AS RELATIVIZED ADJACENCY

Long-distance nasal agreement in Kikongo (Bantu, Congo)

/ k u - d u m u k - i s - i l a /

[k u - d u m u k - i s - i **n** a]

‘to cause to jump for’

(Piggot, 1996)

LOCALITY AS RELATIVIZED ADJACENCY

Long-distance nasal agreement in Kikongo (Bantu, Congo)

/ k u - d u **m** u k - i s - i **l** a /

[k u - d u m u k - i s - i **n** a]

‘to cause to jump for’

- The window is on a ‘nasal/lateral tier’ (and $k=2$).
- /l/ becomes [n] if its neighbor to the left on the tier is a nasal.
- Let’s call such processes **Strictly Tier-Local**.

(Piggot, 1996; Goldsmith, 1979; Lambert, 2023, and others)

QUESTIONS ABOUT THE NATURE OF SUCH PROCESSES

- 1 What is the typological coverage of Strictly Local and Strictly Tier-Local processes?
- 2 Are all regular processes Strictly Local or Strictly Tier-Local?
- 3 Given a phonological process, how can we tell whether it is Strictly Local or Strictly Tier-Local or not?
- 4 To what extent can Strictly Local and Strictly Tier-Local processes be learned?

TYPOLOGICAL COVERAGE OF STRICTLY LOCAL PROCESSES

Chandlee (2014) formally defined strictly local processes and showed:

- 1 Simultaneous application of SPE-style rules
 $A \longrightarrow B / C \text{ --- } D$ for which CAD has a maximum length
are strictly local.
- 2 Approximately 95% of the individual processes in P-Base
(v.1.95, Mielke 2008), including local substitution, deletion,
epenthesis, and synchronic metathesis are strictly local.
- 3 Many *opaque* transformations without any special
treatment are also strictly local

(Chandlee, 2014; Chandlee and Heinz, 2018; Chandlee *et al.*, 2018)

TYPOLOGICAL COVERAGE OF STRICTLY TIER-LOCAL PROCESSES

Can express:

- 1 Self-feeding, iterative spreading (such as nasal spreading in Johore Malay)
- 2 Progressive and regressive consonant and vowel harmonies
 - with and without transparent, opaque segments
 - with and without icy targets
 - with double triggers or positionally sensitive triggers

Cannot express:

- 1 two-sided unbounded contexts
- 2 bi-directional application
- 3 multiple, interacting tiers

(McMullin, 2016; Burness *et al.*, 2021; Lambert and Heinz, 2024)

CHARACTERIZING STRICTLY LOCAL TRANSFORMATIONS

Multiple Equivalent Formalizations of “Strictly Local”

Theory	Result
Logic	?
Automata	?
Algebra	?

Kleene (1956); Büchi (1960); Pin (2021)

Logic

Quantifier Free:

Formulas defining the process can refer to **next** and **previous** string positions, but **cannot use the quantifiers $\exists$ and $\forall$** on the right hand side of $:=$.

Example: Finnish e-raising:

$$\text{high}_{\text{output}}(x) := \\ \text{high}(x) \vee (\text{front}(x) \wedge \neg \text{low}(x) \wedge \text{boundary}(\text{next}(x)))$$

Nesting gives larger windows: **next(next(x))** and so on.

(Lindell and Chandlee, 2016)

LOGICAL CHARACTERIZATIONS OF LOCAL PROCESSES

Logic (continued). Compare:

- 1 $P_{output}(x) := Q(x) \wedge (\exists y)[R(y)]$ (First Order Definable)

Requires scanning whole structure for such a y !!

- 2 $P_{output}(x) := Q(x) \wedge R(x)$ (Quantifier-free Definable)

Information to decide P is **local** to x in the input!!

RESTRICTIVENESS

Fact

- 1 All strictly local and tier-local patterns are regular.
- 2 Not all regular patterns are strictly local or tier-local.

Monadic Second Order Logic $\sim$ **Regular**

|
First Order Logic

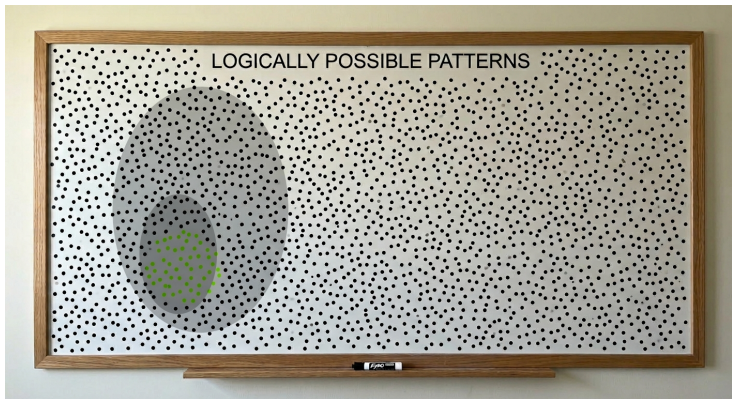
|
Propositional Logic (Quantifier Free)

|
Conjunctions

RESTRICTIVENESS

Fact

- 1 All strictly local and tier-local patterns are regular.
- 2 Not all regular patterns are strictly local or tier-local.



OTHER CHARACTERIZATIONS

For processes over strings, Strictly Local and Strictly Tier-Local processes have other characterizations.

Multiple Equivalent Formalizations of “Strictly Local”

Theory	Result
Logic	Quantifier Free with next/previous
Automata	Finite-state, states $\sim$ last $k-1$ symbols
Algebra	Finite Semigroups, satisfy $xy^\omega = y^\omega$

These other formalizations are useful for **learning** and **verification**.

(Chandlee, 2014; Chandlee *et al.*, 2014; Lindell and Chandlee, 2016; Lambert, 2022, 2023, 2024)

OTHER CHARACTERIZATIONS

For processes over strings, Strictly Local and Strictly Tier-Local processes have other characterizations.

Multiple Equivalent Formalizations of “Strictly Tier Local”

Theory	Result
Logic	Quantifier Free with next/previous on the tier
Automata	Finite-state, states $\sim$ last $k-1$ symbols on tier
Algebra	Finite Semigroups, satisfy $zxy^\omega = zy^\omega$

These other formalizations are useful for [learning](#) and [verification](#).

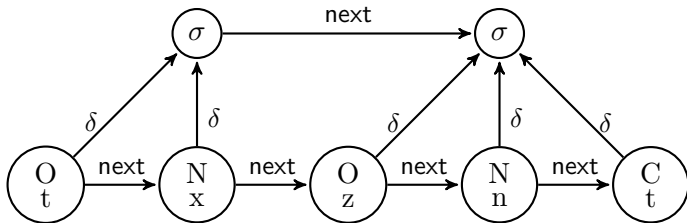
(Chandlee, 2014; Chandlee *et al.*, 2014; Lindell and Chandlee, 2016; Lambert, 2022, 2023, 2024)

LOGIC WITH NON-LINEAR REPRESENTATIONS

Strother-Garcia (2018) shows that the process of syllabification in Imdlawn Tashlhiyt Berber (Dell and Elmedlaoui 1985, Prince and Smolensky 1993) is Quantifier Free.

/ t x z n t /

maps to



LOGIC WITH NON-LINEAR REPRESENTATIONS

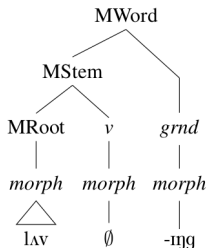
Strother-Garcia (2018) concludes “...syllabification in ITB can be represented by a QF [logical] transduction, a formalism restricted to *substantially lower computational complexity* than [traditional] phonological grammars...Establishing that ITB syllabification is QF highlights *an insight not apparent* from [traditional] grammatical formalisms...”

LOGIC WITH NON-LINEAR REPRESENTATIONS

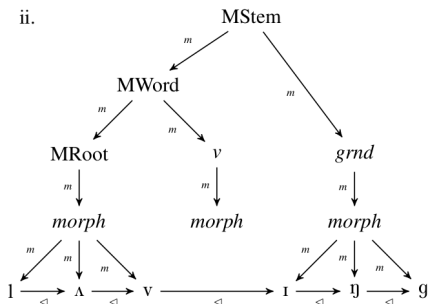
Dolatian (2020) examines the phonology-morphology interface in Armenian in light of Quantifier Free logical transductions. He concludes

*“the bulk of the morphology-phonology interface requires **local computation**, not global computation.”*

i.



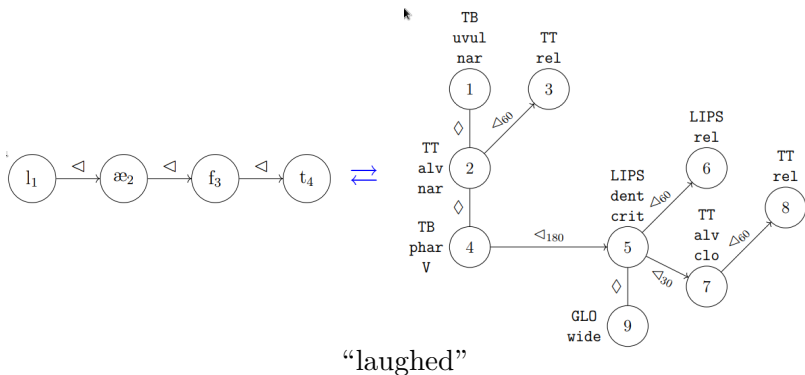
ii.



“loving”

LOGIC WITH NON-LINEAR REPRESENTATIONS

Nelson (2022, 2024) examines the phonetics-phonology interface, and shows how to convert coupling graphs in Articulatory Phonology to familiar segmental representations **and vice versa** using First Order logic. Whether it is QF remains open.



WHAT'S NEXT

- 1 What 'Regular' Means
- 2 Locality/Subregularity in Phonological Processes
- 3 **Locality/Subregularity in Phonological Constraints**
- 4 Learning Phonological Constraints
- 5 Learning Phonological Processes
- 6 Beyond Phonology
- 7 Conclusion

ARE POSSIBLE MARKEDNESS CONSTRAINTS COMPUTATIONALLY BOUNDED?

Finnish

- /#reke#/ → [#reki#]
- Constraint: *e#

Kikongo

- /ku-dumuk-is-ila/ → [ku-dumuk-is-ina]
- Constraint: *[+nasal]...[+lateral]

Containment

**Markedness constraints are ill-formed structures...
and well-formed structures do not contain them!**

(Rogers, 1998; Heinz, 2010; Rogers *et al.*, 2013; Rogers and Lambert, 2019; Lambert *et al.*, 2021)

A RESTRICTED LOGICAL LANGUAGE

Grammars are Conjunctions of negative literals

$$\neg C_1 \wedge \neg C_2 \dots \wedge \neg C_n$$

Plain English

“A well-formed structure
must not contain C_1 and
must not contain C_2 and
...
must not contain C_n ”

Constraints C_i

The constraints C_i are structures connected by relations such as **next**, **next-on-tier** or others.

A RESTRICTED LOGICAL LANGUAGE

Grammars are Conjunctions of negative literals

$$\neg C_1 \wedge \neg C_2 \dots \wedge \neg C_n$$

Monadic Second Order Logic $\sim$ **Regular**

|

First Order Logic

|

Propositional Logic (Quantifier Free)

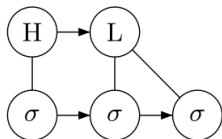
|

Conjunctions

EXAMPLE: AUTOSEGMENTAL CONSTRAINTS



félàlà HLL ‘junction’
(Mende)



Some well-studied patterns of tonal association

- 1 Position-specific contours (Hausa)
- 2 Position-specific plateaus (Northern Karanga)
- 3 Melody constraints (Mende)
- 4 Quality-dependent plateaus (Kukuya)

Jardine (2016, 2017) argues grammars like

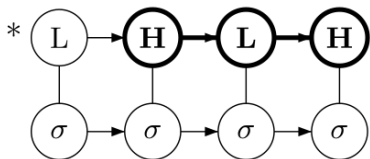
$$\neg C_1 \wedge \neg C_2 \dots \wedge \neg C_n$$

where C_i are autosegmental structures captures the typology better than previous OT and derivational analyses.

Mende tonotactic grammar

$$\neg HLH$$

$$HLH = \textcircled{H} \rightarrow \textcircled{L} \rightarrow \textcircled{H}$$

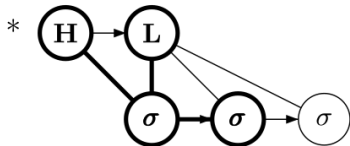
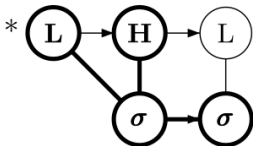
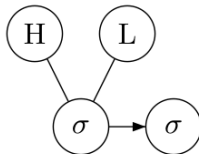


(Jardine, 2016, 2017)

Mende tonotactic grammar

$$\neg HLH \wedge \neg \text{NonFinal-Contour}$$

NonFinal-Contour =



(Jardine, 2016, 2017)

SUMMARY: AUTOSEGMENTAL CONSTRAINTS

Some well-studied patterns of tonal association

- 1 Position-specific contours
- 2 Position-specific plateaus
- 3 Melody constraints
- 4 Quality-dependent plateaus

Jardine (2016, 2017) argues grammars like

$$\neg C_1 \wedge \neg C_2 \dots \wedge \neg C_n$$

where C_i are autosegmental structures captures the typology better than previous OT and derivational analyses.

TYPOLOGICAL COVERAGE

Lambert (2026) studies **phonotactic patterns** over strings.

*“In this article, I classify a **wide range of attested long-distance phonological markedness constraints**, covering **stress**, **harmony** and **tone**, with focus on interactions of constraints over multiple tiers.*

*To this end, connections are established between **propositional logic over multiple tiers** and **algebraic properties of formal languages**. These techniques allow for **mechanical verification or refutation** of membership in a class...*

*I demonstrate that the **entire range lies within the propositional level**, including Uyghur backness harmony and Karanga Shona tone, which have been presented as challenges to aspects of the subregular hypothesis.” (Lambert, 2026)*

A MORE EXPRESSIVE LOGICAL LANGUAGE

Grammars are a Propositional Logic of literals

$$C_1 \wedge \neg C_2 \dots \wedge (C_{n-1} \vee C_n)$$

Plain English

“A well-formed structure
must contain C_1 and
must not contain C_2 and
...
must contain either C_{n-1} or
 C_n ”

Constraints C_i

The constraints C_i are structures connected by relations such as **next**, **next-on-tier** or **precedes**.

A MORE EXPRESSIVE LOGICAL LANGUAGE

Grammars are a Propositional Logic of literals

$$C_1 \wedge \neg C_2 \dots \wedge (C_{n-1} \vee C_n)$$

Monadic Second Order Logic $\sim$ **Regular**

|

First Order Logic

|

Propositional Logic (Quantifier Free)

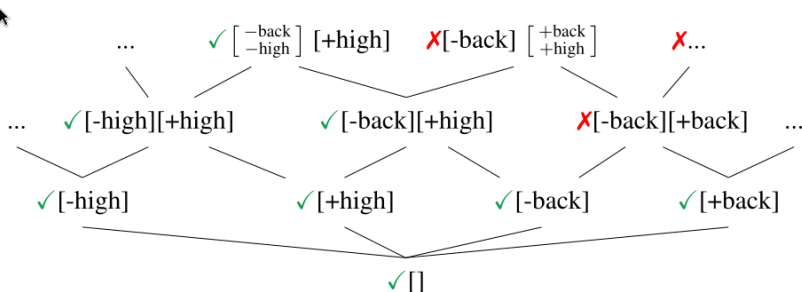
|

Conjunctions

WHAT'S NEXT

- 1 What 'Regular' Means
- 2 Locality/Subregularity in Phonological Processes
- 3 Locality/Subregularity in Phonological Constraints
- 4 **Learning Phonological Constraints**
- 5 Learning Phonological Processes
- 6 Beyond Phonology
- 7 Conclusion

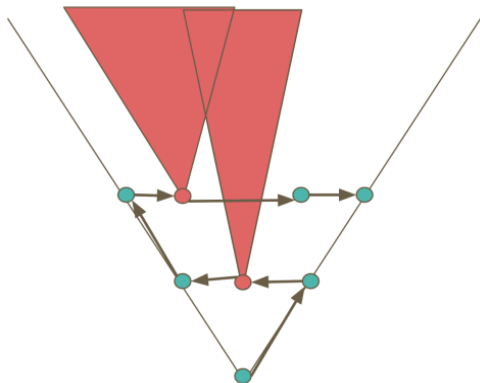
CONSTRAINT SPACE IS STRUCTURED BY CONTAINMENT



Possible constraints are partially ordered with **superfactors** containing their **subfactors**. **Illicit** structures are upward entailing, **licit** structures are downward entailing.

(Chandlee *et al.*, 2019; Rawski, 2021; Swanson *et al.*, 2026)

BUFIA: BOTTOM-UP FACTOR INFERENCE ALGORITHM



The BUFIA algorithm traverses and prunes the search space. Arrows indicate direction of search, while shaded areas indicate sections of the partial order which have been pruned. Chandlee et al (2019) prove the most general constraints true of the data are found.

(Chandlee *et al.*, 2019; Rawski, 2021; Swanson *et al.*, 2026)

BUFIA

- ① is **deterministic** and returns grammars of the form

$$\neg C_1 \wedge \neg C_2 \dots \wedge \neg C_n$$

- ② Words are well-formed if they **satisfy** the grammar.
- ③ Can be implemented over different relational structures: strings, feature-bundle strings, autosegmental representations, trees, ...

Maximum Entropy based UCLA Phonotactic Learner

- ① is **probabilistic** and returns weighted constraints.
- ② Words are well-formed according to the **weighted sum** of their constraint violation profiles.
- ③ Unclear how to implement for representations other than strings.

(Chandlee *et al.*, 2019; Hayes and Wilson, 2008)

COMPARING BUFIA WITH STATISTICAL MODELS

① South Bolivian Quechua Phonotactics:

BUFIA-FTR $\sim$ MAXENT-FTR $>$ BUFIA-SEG,
MAXENT-SEG

(Wilson and Gallagher, 2018; Swanson *et al.*, 2026)

COMPARING BUFIA WITH STATISTICAL MODELS

① South Bolivian Quechua Phonotactics:

BUFIA-FTR $\sim$ MAXENT-FTR $>$ BUFIA-SEG,
MAXENT-SEG

(Wilson and Gallagher, 2018; Swanson *et al.*, 2026)

② Stress in Jordanian, Egyptian, and Sudanese Arabic:

BUFIA $>$ Neural Encoder-Decoder **on less data** but Neural
Encoder-Decoder $>$ BUFIA **on more data**

(Qaddoumi *et al.*, 2026)

COMPARING BUFIA WITH STATISTICAL MODELS

- 1 South Bolivian Quechua Phonotactics:
BUFIA-FTR $\sim$ MAXENT-FTR > BUFIA-SEG,
MAXENT-SEG
(Wilson and Gallagher, 2018; Swanson *et al.*, 2026)
- 2 Stress in Jordanian, Egyptian, and Sudanese Arabic:
BUFIA > Neural Encoder-Decoder on less data but Neural
Encoder-Decoder > BUFIA on more data
(Qaddoumi *et al.*, 2026)
- 3 Modeling Bilingual Phonotactic Judgments:
BUFIA-FTR > MAXENT-FTR
(Tetzloff and Li, 2026)

COMPARING BUFIA WITH STATISTICAL MODELS

- 1 South Bolivian Quechua Phonotactics:
BUFIA-FTR $\sim$ MAXENT-FTR $>$ BUFIA-SEG,
MAXENT-SEG
(Wilson and Gallagher, 2018; Swanson *et al.*, 2026)
- 2 Stress in Jordanian, Egyptian, and Sudanese Arabic:
BUFIA $>$ Neural Encoder-Decoder on less data but Neural
Encoder-Decoder $>$ BUFIA on more data
(Qaddoumi *et al.*, 2026)
- 3 Modeling Bilingual Phonotactic Judgments:
BUFIA-FTR $>$ MAXENT-FTR
(Tetzloff and Li, 2026)
- 4 Hausa Tonalactics:
BUFIA-AR $>$ BUFIA-SEG, BUFIA-FTR,
MAXENT-FTR, MAXENT-SEG
(Li and Heinz, 2026)

LEARNED TONOTACTIC CONSTRAINTS IN HAUSA

	$\mu = 1$	$\mu = 2$	$\mu = 3$
$s = 1$	$\begin{array}{c} L \quad H \\ \diagdown \quad \diagup \\ \mu \\ \\ \sigma \end{array}$ (a) $\begin{array}{c} H \quad L \\ \diagdown \quad \diagup \\ \mu \\ \\ \sigma \end{array}$ (b)	$\begin{array}{c} L \quad H \\ \quad \\ \mu \quad \mu \\ \diagdown \quad \diagup \\ \sigma \end{array}$ (c)	$\begin{array}{c} \mu \quad \mu \quad \mu \\ \diagdown \quad \quad \diagup \\ \sigma \end{array}$ (d)
$s = 2$	-	-	$\begin{array}{c} H \quad L \\ \quad \diagup \\ \mu \quad \mu \\ \quad \\ \sigma \quad \sigma \end{array}$ (e) $\begin{array}{c} H \quad L \\ \diagdown \quad \\ \mu \quad \mu \\ \quad \diagdown \\ \sigma \quad \sigma \end{array}$ (f) $\begin{array}{c} H \quad L \quad H \\ \diagup \quad \diagdown \quad \\ \mu \quad \mu \quad \mu \\ \diagdown \quad \\ \sigma \quad \sigma \end{array}$ (g)
$s = 3$	-	-	$\begin{array}{c} H \quad L \\ \quad \diagup \\ \mu \quad \mu \\ \quad \\ \sigma \quad \sigma \end{array}$ (h) $\begin{array}{c} L \quad H \quad L \\ \diagdown \quad \diagdown \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ \sigma \quad \sigma \quad \sigma \end{array}$ (i) $\begin{array}{c} H \quad L \quad H \\ \diagdown \quad \\ \mu \quad \mu \\ \quad \quad \\ \sigma \quad \sigma \quad \sigma \end{array}$ (j) $\begin{array}{c} L \quad H \quad L \\ \quad \diagdown \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ \sigma \quad \sigma \quad \sigma \end{array}$ (k)

Table 9: Hausa constraints returned by BUFIA-AR when $s \leq 3$, $t \leq 3$, $\mu \leq 3$

(Jardine, 2017; Li and Heinz, 2026, to appear, in *Phonology*)

A SECOND THEME

optimization $\sim$ **satisfaction**

statistics $\sim$ **structure**

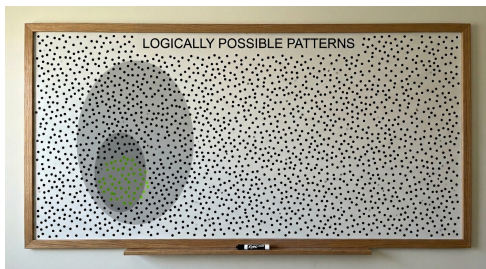


WHAT'S NEXT

- 1 What 'Regular' Means
- 2 Locality/Subregularity in Phonological Processes
- 3 Locality/Subregularity in Phonological Constraints
- 4 Learning Phonological Constraints
- 5 **Learning Phonological Processes**
- 6 Beyond Phonology
- 7 Conclusion

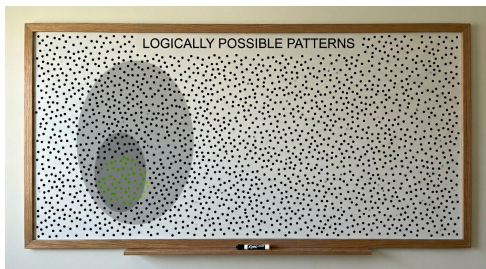
LEARNING TRANSFORMATIONS

- 1 There are **theoretical results** for learning subregular classes such as **Strictly Local** and **Tier Local Processes** ...
(Chandlee *et al.*, 2014; Jardine *et al.*, 2014; Burness and McMullin, 2021)



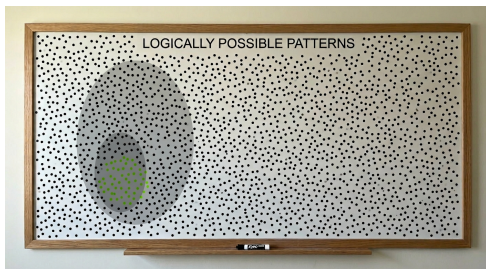
LEARNING TRANSFORMATIONS

- 1 There are **theoretical results** for learning subregular classes such as **Strictly Local** and **Tier Local Processes** ...
(Chandlee *et al.*, 2014; Jardine *et al.*, 2014; Burness and McMullin, 2021)
- 2 ...in contrast to **the full regular class** for which theoretical results are **uniformly negative**
(Angluin, 1980; Valiant, 1984; de la Higuera, 2010)



LEARNING TRANSFORMATIONS

- 1 There are **theoretical results** for learning subregular classes such as **Strictly Local** and **Tier Local Processes** ...
(Chandlee *et al.*, 2014; Jardine *et al.*, 2014; Burness and McMullin, 2021)
- 2 ...in contrast to **the full regular class** for which theoretical results are **uniformly negative**
(Angluin, 1980; Valiant, 1984; de la Higuera, 2010)
- 3 If there are too many ways to interpret the data, learning is hard! **Structure helps!** (Rawski and Heinz, 2019)



LINGUISTIC REPRESENTATIONS MATTER TOO

- 1 Markowska (2026) shows that **phonological features dramatically reduces** the number of examples sufficient to generalize correctly in a variety of case studies.

(Markowska, 2026; Markowska and Heinz, 2026)

LINGUISTIC REPRESENTATIONS MATTER TOO

- 1 Markowska (2026) shows that **phonological features dramatically reduces** the number of examples sufficient to generalize correctly in a variety of case studies.
- 2 Her methods synthesize BUFIA with earlier work to learn the behavior of each feature in a process separately.
- 3 This reduces the number of examples sufficient **from 1,000s or 10,000s to hundreds or dozens.**

(Markowska, 2026; Markowska and Heinz, 2026)

LINGUISTIC REPRESENTATIONS MATTER TOO

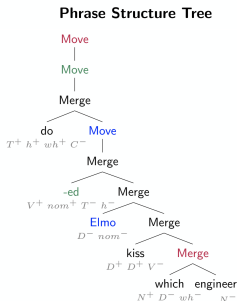
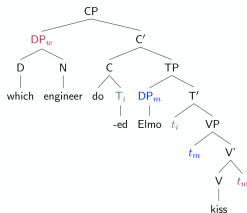
- 1 Markowska (2026) shows that **phonological features dramatically reduces** the number of examples sufficient to generalize correctly in a variety of case studies.
- 2 Her methods synthesize BUFIA with earlier work to learn the behavior of each feature in a process separately.
- 3 This reduces the number of examples sufficient **from 1,000s or 10,000s to hundreds or dozens**.
- 4 For example, she shows that one can learn **the opaque interaction between ɔ-raising and word-final devoicing in Polish** with as few as 300 forms sampled from the most frequent words in child-directed speech, *of which only ~3% show any alternation!*

(Markowska, 2026; Markowska and Heinz, 2026)

WHAT'S NEXT

- 1 What 'Regular' Means
- 2 Locality/Subregularity in Phonological Processes
- 3 Locality/Subregularity in Phonological Constraints
- 4 Learning Phonological Constraints
- 5 Learning Phonological Processes
- 6 **Beyond Phonology**
- 7 Conclusion

SYNTAX



- Rogers (1997, 1998) showed that Context-Free Languages are Strictly Local Tree Structures.
- Graf (2022) and colleagues argue many syntactic constraints are Strictly Local or Tier Local either directly over derivation trees or over sequential paths present in those trees (see also Hanson, 2023, 2025)
- Swanson (2025) shows how a BUFA-style algorithm can learn such tree-constraints from example derivation trees.

- Pietroski (2018b,a) argues for limiting semantic types and **using conjunction as the principal mode of semantic composition.**
- Pietroski and Icard (2025) show this restricted logic retains the **computational complexity of propositional logic.**

COGNITIVE PARALLELISM

Computationally Possible

Context-Sensitive

Regular

S

P

Subregular

trees

strings

WHAT'S NEXT

- 1 What 'Regular' Means
- 2 Locality/Subregularity in Phonological Processes
- 3 Locality/Subregularity in Phonological Constraints
- 4 Learning Phonological Constraints
- 5 Learning Phonological Processes
- 6 Beyond Phonology
- 7 **Conclusion**

SOME UNTURNED STONES, PART 1

Symbolic, formal systems provide insight into natural languages but are still underexplored.

- 1 **Fine-grained computational properties** as necessary linguistic properties!
- 2 **Linguistic representations** matter too!
- 3 **Logic** provides a way to talk about both simultaneously.

SOME UNTURNED STONES, PART 2

Learning as Satisficing, not learning as optimizing!
(cf. BUFIA, Yolyan (2025); Belth (2026))

- 1 Start with a simple grammar and test it against the data.
- 2 Complexify in specific ways—based on representation and computational class—and test again.
- 3 Repeat until a “sufficiently good” hypothesis is found.

MORE STONES TO TURN

- 1 More careful analysis of complex cases: bidirectionality, interaction between tiers, interacting processes
- 2 Decomposing observed forms into lexicons (URs), morphological and phonological grammars
- 3 Grammars with exceptions
- 4 Modeling course of acquisition
- 5 Study of more logical languages
- 6 Using these properties to understand other types of computation

CONCLUSION

- 1 Don't underestimate the power of **structure** and **satisfaction**!
- 2 Mathematical and computational methods based in formal grammars and symbolic computation for linguistics are **both viable and nascent**.
- 3 There is much work to do!



ACKNOWLEDGMENTS

Many thanks to Thomas Graf, Han Li, Magda Markowska, Scott Nelson, William Oliver, Owen Rambow, Jon Rawski, and the Stony Brook Linguistics Brown Bag lunch audience for helpful feedback and discussion.

Thank you!

This work is supported by NSF BCS #2416183.

REFERENCES I

- Angluin, Dana. 1980. Inductive inference of formal languages from positive data. *Information Control* 45:117–135.
- Baković, Eric. 2000. Harmony, dominance and control. Doctoral dissertation, Rutgers University.
- Belth, Caleb. 2026. A learning-based account of phonological tiers. *Linguistic Inquiry* 57:229–265.
- Büchi, J. Richard. 1960. Weak second-order arithmetic and finite automata. *Mathematical Logic Quarterly* 6:66–92.
- Burness, Phillip, and Kevin McMullin. 2021. Learning multiple independent tier-based processes. In *Proceedings of the Fifteenth International Conference on Grammatical Inference*, edited by Jane Chandlee, Rémi Eyraud, Jeff Heinz, Adam Jardine, and Menno van Zaanen, vol. 153 of *Proceedings of Machine Learning Research*, 66–80. PMLR.

REFERENCES II

URL

<https://proceedings.mlr.press/v153/burness21a.html>

Burness, Phillip A., Kevin J. McMullin, and Jane Chandlee.

2021. Long-distance phonological processes as tier-based strictly local functions. *Glossa: a Journal of General Linguistics* 6.

Chandlee, Jane. 2014. Strictly local phonological processes.

Doctoral dissertation, The University of Delaware.

Chandlee, Jane, Rémi Eyraud, and Jeffrey Heinz. 2014.

Learning strictly local subsequential functions. *Transactions of the Association for Computational Linguistics* 2:491–503.

REFERENCES III

- Chandlee, Jane, Remi Eyraud, Jeffrey Heinz, Adam Jardine, and Jonathan Rawski. 2019. Learning with partially ordered representations. In *Proceedings of the 16th Meeting on the Mathematics of Language*, 91–101. Toronto, Canada: Association for Computational Linguistics.
- Chandlee, Jane, and Jeffrey Heinz. 2018. Strict locality and phonological maps. *Linguistic Inquiry* 49:23–60.
- Chandlee, Jane, Jeffrey Heinz, and Adam Jardine. 2018. Input strictly local opaque maps. *Phonology* 35:171–205.
- Chandlee, Jane, and Adam Jardine. 2019. Autosegmental input-strictly local functions. *Transactions of the Association for Computational Linguistics* 7:157–168.
- Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 113–124. IT-2.

REFERENCES IV

- Dolatian, Hossep. 2020. Computational locality of cyclic phonology in Armenian. Doctoral dissertation, Stony Brook University.
- Finley, Sara. 2008. The formal and cognitive restrictions on vowel harmony. Doctoral dissertation, Johns Hopkins University, Baltimore, MD.
- Finley, Sara. 2011. The privileged status of locality in consonant harmony. *Journal of Memory and Language* 65:74–83.
- Goldsmith, John. 1979. *Autosegmental phonology*. New York: Garland.

REFERENCES V

- Graf, Thomas. 2012. Locality and the complexity of Minimalist derivation tree languages. In *Formal Grammar 2010/2011*, edited by Philippe de Groot and Mark-Jan Nederhof, vol. 7395 of *Lecture Notes in Computer Science*, 208–227. Heidelberg: Springer.
URL http://dx.doi.org/10.1007/978-3-642-32024-8_14
- Graf, Thomas. 2022. Subregular linguistics: bridging theoretical linguistics and formal grammar. *Theoretical Linguistics* 48:145–184.
- Hanson, Kenneth. 2023. A TSL analysis of Japanese case. In *Proceedings of the Society for Computation in Linguistics 2023*, edited by Tim Hunter and Brandon Prickett, 15–24. Amherst, MA: Association for Computational Linguistics.
URL <https://aclanthology.org/2023.scil-1.2/>

REFERENCES VI

- Hanson, Kenneth. 2025. Tier-based strict locality and the typology of agreement. *Journal of Language Modelling* 13.
URL <https://doi.org/10.15398/jlm.v13i1.411>
- Hao, Sophie. 2024. Universal Generation for Optimality Theory Is PSPACE-Complete. *Computational Linguistics* 1–35.
- Hao, Yiding. 2019. Finite-state optimality theory: non-rationality of harmonic serialism. *Journal of Language Modelling* 7:49–99.
- Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:379–440.
- Heinz, Jeffrey. 2010. Learning long-distance phonotactics. *Linguistic Inquiry* 41:623–661.

REFERENCES VII

- Heinz, Jeffrey. 2018. The computational nature of phonological generalizations. In *Phonological Typology*, edited by Larry Hyman and Frans Plank, Phonetics and Phonology, chap. 5, 126–195. De Gruyter Mouton.
- Heinz, Jeffrey. 2025. The logical structure of phonological generalizations. *Phonological Studies* 28:69–80.
- Heinz, Jeffrey, and Regine Lai. 2013. Vowel harmony and subsequentiality. In *Proceedings of the 13th Meeting on the Mathematics of Language (MoL 13)*, edited by Andras Kornai and Marco Kuhlmann, 52–63. Sofia, Bulgaria.
- de la Higuera, Colin. 2010. *Grammatical Inference: Learning Automata and Grammars*. Cambridge University Press.

REFERENCES VIII

- Huybregts, Riny. 1984. The weak inadequacy of context-free phrase structure grammars. In *Van periferie naar kern*, edited by Ger de Haan, Mieke Trommelen, and Wim Zonneveld, 81–99. Dordrecht, The Netherlands: Foris.
- Jardine, Adam. 2016. Locality and non-linear representations in tonal phonology. Doctoral dissertation, University of Delaware.
- Jardine, Adam. 2017. The local nature of tone-association patterns. *Phonology* 34:385–405.
- Jardine, Adam, Jane Chandlee, Rémi Eyraud, and Jeffrey Heinz. 2014. Very efficient learning of structured classes of subsequential functions from positive data. In *Proceedings of the Twelfth International Conference on Grammatical Inference (ICGI 2014)*, edited by Alexander Clark, Makoto Kanazawa, and Ryo Yoshinaka, vol. 34, 94–108. JMLR: Workshop and Conference Proceedings.

REFERENCES IX

- Johnson, C. Douglas. 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.
- Kaplan, Ronald, and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics* 20:331–378.
- Kleene, S.C. 1956. Representation of events in nerve nets. In *Automata Studies*, edited by C.E. Shannon and J. McCarthy, 3–40. Princeton University. Press.
- Kobele, Gregory M. 2011. Minimalist tree languages are closed under intersection with recognizable tree languages. In *Logical Aspects of Computational Linguistics*, edited by Sylvain Pogodalla and Jean-Philippe Prost, 129–144. Berlin, Heidelberg: Springer Berlin Heidelberg.

REFERENCES X

Lambert, Dakotah. 2022. Unifying classification schemes for languages and processes with attention to locality and relativizations thereof. Doctoral dissertation, Stony Brook University.

URL

<https://vvulpes0.github.io/PDF/dissertation.pdf/>

Lambert, Dakotah. 2023. Relativized adjacency. *Journal of Logic Language and Information* 32:707–731.

Lambert, Dakotah. 2024. System description: A theorem-prover for subregular systems: The language toolkit and its interpreter, plebby. In *Functional and Logic Programming*, edited by Jeremy Gibbons and Dale Miller, 311–328. Singapore: Springer Nature Singapore.

Lambert, Dakotah. 2026. Multitier phonotactics with logic and algebra. *Phonology* 43:1–31.

REFERENCES XI

- Lambert, Dakotah, and Jeffrey Heinz. 2024. Algebraic reanalysis of phonological processes described as output-oriented. In *Proceedings of the Society for Computation in Linguistics*, vol. 7, 129–138.
- Lambert, Dakotah, Jonathan Rawski, and Jeffrey Heinz. 2021. Typology emerges from simplicity in representations and learning. *Journal of Language Modelling* 9:151–194.
- Lamont, Andrew. 2021. Optimizing over subsequences generates context-sensitive languages. *Transactions of the Association for Computational Linguistics* 9:528–537.
- Lamont, Andrew. 2022. Optimality theory implements complex functions with simple constraints. *Phonology* 38:729–740.

REFERENCES XII

- Li, Han, and Jeffrey Heinz. 2026. What matters in tonotactic learning. In *Proceedings of the Society for Computation in Linguistics 2026*, edited by Rob Voigt, Alex Warstadt, Naomi Feldman, and Tal Linzen, 180–190. San Diego, CA: Association for Computational Linguistics.
URL <https://aclanthology.org/2026.scil-main.18/>
- Li, Han, and Jeffrey Heinz. to appear. Tonotactic pattern learning over autosegmental representations. *Phonology* .
- Lindell, Steven, and Jane Chandlee. 2016. A logical characterization of input strictly local functions Presented at the Fourth Workshop on Natural Language and Computer Science, in affiliation with LICS 2016.

REFERENCES XIII

- Lombardi, Lindsa. 1999. Positional faithfulness and voicing assimilation in optimality theory. *Natural Language and Linguistic Theory* 17:267–302.
URL <https://doi.org/10.1023/A:1006182130229>
- Markowska, Magda. 2026. Learning phonological transductions from sparse data over structured representations. Doctoral dissertation, Stony Brook University.
- Markowska, Magdalena, and Jeffrey Heinz. 2026. Learning k-isl functions with phonological features. *Journal of Language Modelling* 14:1–80.
URL <https://jlm.ipipan.waw.pl/index.php/JLM/article/view/493>
- McMullin, Kevin. 2016. Tier-based locality in long-distance phonotactics: Learnability and typology. Doctoral dissertation, University of British Columbia.

REFERENCES XIV

- Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. Oxford: Oxford University Press.
- Nelson, Scott. 2022. Are representations in articulatory and generative phonology so different? Poster presented at the 2022 Annual Meeting on Phonology.
- Nelson, Scott. 2024. The computational structure of phonological and phonetic knowledge. Doctoral dissertation, Stony Brook University.
- Odden, David. 2013. *Introducing Phonology*. 2nd ed. Cambridge University Press.
- Pietroski, Paul M. 2018a. *Conjoining Meanings*. Oxford University Press.

REFERENCES XV

- Pietroski, Paul M. 2018b. Limiting semantic types. In *Language, Syntax, and the Natural Sciences*, edited by Angel Gallego and Roger Martin, 192–211. Cambridge University Press.
- Pietroski, Paul M., and Thomas Icard. 2025. Simple concepts and meanings: Adequacy from below. *Meaning* 1:1–39.
- Piggot, Glynn. 1996. Implications of consonant nasalization for a theory of harmony. *Canadian Journal of Linguistics* 41.
- Pin, Jean-Éric. 2021. *Handbook of Automata Theory*. European Mathematical Society Publishing House.
URL <https://hal.archives-ouvertes.fr/hal-03579131>

REFERENCES XVI

- Qaddoumi, Abed, Salam Khalifa, Jordan Kodner, Jeffrey Heinz, and Owen Rambow. 2026. Learning stress in Arabic low-resource settings. In *Proceedings of the Society for Computation in Linguistics 2026*, edited by Rob Voigt, Alex Warstadt, Naomi Feldman, and Tal Linzen, 262–279. San Diego, CA: Association for Computational Linguistics. URL <https://aclanthology.org/2026.scil-main.24/>
- Rawski, Jonathan. 2021. Structure and learning in natural language. Doctoral dissertation, Stony Brook University.
- Rawski, Jonathan, and Jeffrey Heinz. 2019. No free lunch in linguistics or machine learning: Response to Pater. *Language* 95:e125–e135.
- Riggle, Jason. 2004. Generation, recognition, and learning in finite state Optimality Theory. Doctoral dissertation, University of California, Los Angeles.

REFERENCES XVII

Rogers, James. 1997.

Strict LT_2 : Regular :: Local : Recognizable. In *Logical Aspects of Computational Linguistics: First International Conference, LACL '96 (Selected Papers)*, edited by Christian Retoré, vol. 1328 of *Lectures Notes in Computer Science/Lectures Notes in Artificial Intelligence*, 366–385. Springer.

Rogers, James. 1998. *A Descriptive Approach to Language-Theoretic Complexity*. Stanford, CA: CSLI Publications.

Rogers, James, Jeffrey Heinz, Margaret Fero, Jeremy Hurst, Dakota Lambert, and Sean Wibel. 2013. Cognitive and sub-regular complexity. In *Formal Grammar*, edited by Glyn Morrill and Mark-Jan Nederhof, vol. 8036 of *Lecture Notes in Computer Science*, 90–108. Springer.

REFERENCES XVIII

- Rogers, James, and Dakota Lambert. 2019. Some classes of sets of structures definable without quantifiers. In *Proceedings of the 16th Meeting on the Mathematics of Language*, 63–77. Toronto, Canada: Association for Computational Linguistics. URL <https://www.aclweb.org/anthology/W19-5706>
- Shieber, Stuart. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8:333–343.
- Stabler, Edward P. 2019. Three mathematical foundations for syntax. *Annual Review of Linguistics* 5:243–260.
- Strother-Garcia, Kristina. 2018. Imdlawn Tashlhiyt Berber syllabification is quantifier-free. In *Proceedings of the Society for Computation in Linguistics*, vol. 1. Article 16.
- Strother-Garcia, Kristina. 2019. Using model theory in phonology: A novel characterization of syllable structure and syllabification. Doctoral dissertation, University of Delaware.

REFERENCES XIX

- Swanson, Logan. 2025. Learning multi tier-based strictly 2-local languages. In *Proceedings of the 18th Meeting of the Mathematics of Language*.
- Swanson, Logan, Jeffrey Heinz, and Jonathan Rawski. 2026. Phonotactic learning with structure, not statistics. *Linguistic Inquiry* 1–24.
URL <https://doi.org/10.1162/LING.a.584>
- Tetzloff, Katerina, and Han Li. 2026. Gradient bilingual phonotactic judgments reflect dominance-weighted grammar mixture. Presentation at the Annual Meeting of Phonology, UMD.
- Valiant, L.G. 1984. A theory of the learnable. *Communications of the ACM* 27:1134–1142.

REFERENCES XX

- Wilson, Colin, and Gillian Gallagher. 2018. Accidental gaps and surface-based phonotactic learning: a case study of South Bolivian Quechua. *Linguistic Inquiry* 49:610–623.
- Yolyan, Tatevik. 2025. A framework for learning phonological maps as logical transductions. In *Proceedings of the 18th Meeting on the Mathematics of Language*, edited by Thomas Graf, Matthew Hayden, Jeffrey Heinz, Lawrence Moss, and Yang Wang, 44–58. Stony Brook, NY, USA: Association for Computational Linguistics.
- URL <https://aclanthology.org/2025.mol-1.4/>