Learning with Locality Across Speech and Sign

Jon Rawski
Jane Chandlee

Dept. of Linguistics and Institute for Advanced Computational Science
Stony Brook University
Tri-Co Dept of Linguistics, Haverford College

May 7, 2019

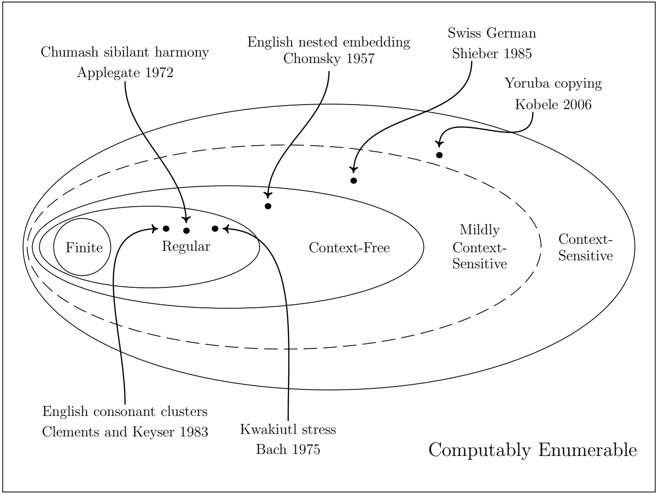
Workshop Goal

- Understand phonological rules across speech and sign
- How do they constrain grammar hypotheses learners infer from data?
- How do learners balance these limits with modality properties?

Computational Perspective

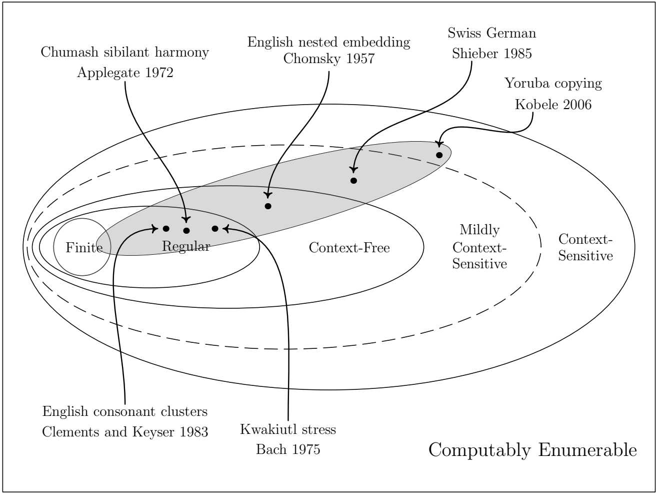
- Mathematically precise representational claims
- Mathematically restricted grammars
- Efficient, provably correct learning algorithms

Learning Depends on Hypothesis Spaces



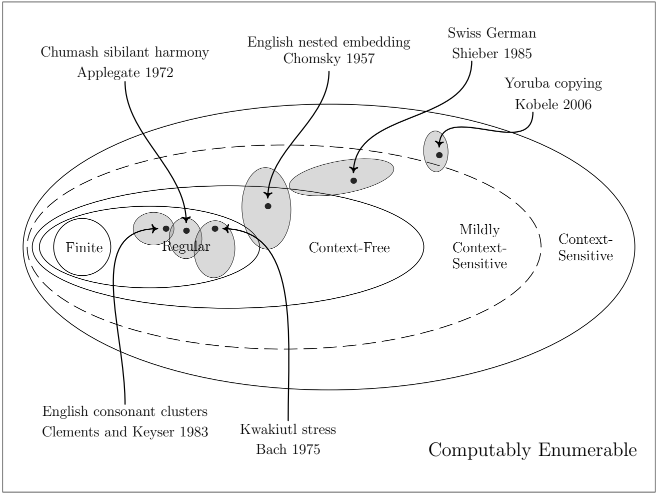
Heinz & Idsardi 2013, Rawski & Heinz 2019

Learning Depends on Hypothesis Spaces



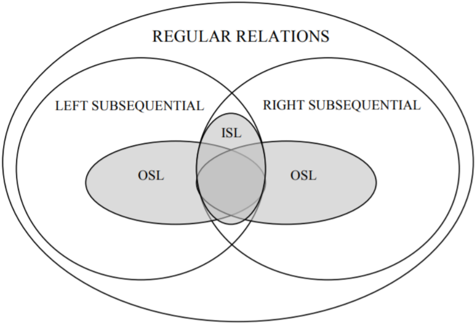
Heinz & Idsardi 2013, Rawski & Heinz 2019

Learning Depends on Hypothesis Spaces



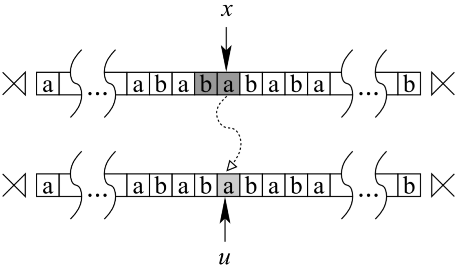
Heinz & Idsardi 2013, Rawski & Heinz 2019

Phonology is Subregular



- Regular/finite-state: Memory does not grow with input size
- Sufficient for phonology (Johnson 1972, Kaplan & Kay 1994)
  - Underlying/Surface pairs: (ba:d, ba:t),
  - Rewrite Rules: -son ! -voice /_#,
  - Constraint Interaction: *[-son,+voice]# >> IDENT(voi)
- New hypothesis: phonology only needs **subregular** power

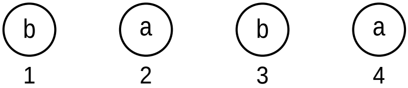
Input Strictly Local Functions



ISL Functions (Chandlee 2014, Chandlee & Heinz 2018)

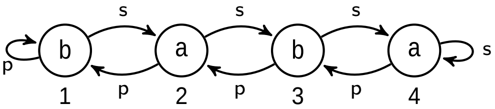
- output *u* tracks contiguous input substrings *x* of length *k*
  - (CAD, CBD), A ! B / C__D, *CAD >> FAITH(A ! B)
  - Intervocalic Voicing is 3-ISL: VTV ! VDV
- About 95% of processes in P-Base (Mielke 2008)
  - Substitution, deletion, epenthesis, general affixation, metathesis (local & bounded nonlocal), partial reduplication, ...

What information is present in a string?



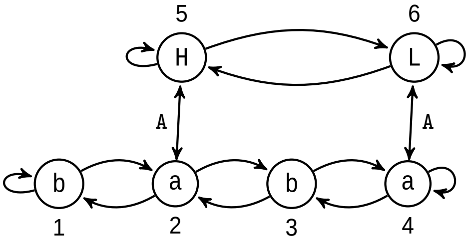
- Domain of sequence elements *f*1,2,3,4*g*
- Labeling relations *f*a,b*g*(IPA, features, orthography, etc)
- Ordering functions (successor and predecessor)

What information is present in a string?



- Domain of sequence elements *f*1,2,3,4*g*
- Labeling relations *f*a,b*g*(IPA, features, orthography, etc)
- Ordering functions (successor and predecessor)

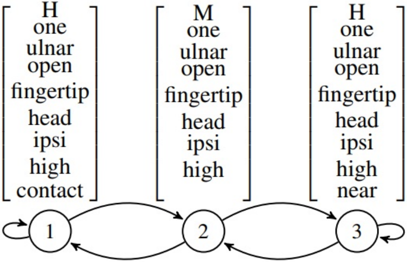
What information is present in a string?



- Domain of sequence elements $\{1,2,3,4,5,6\}$
- Labeling relations $\{$a,b,H,L$\}$ (now including tone)
- Association relations
- Ordering functions (successor and predecessor)

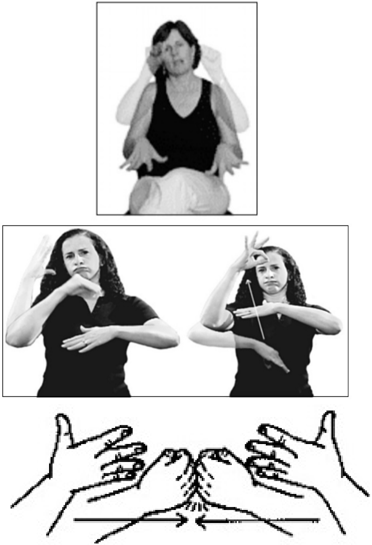
Signs as Strings

IDEA (ASL)

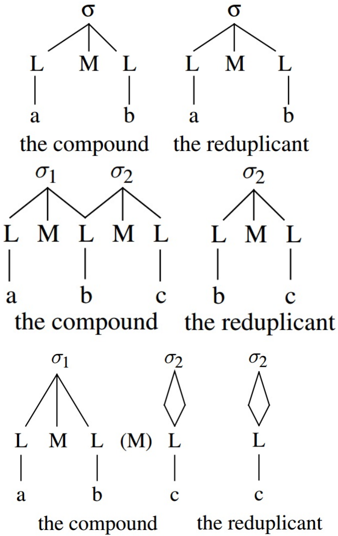


- Models signs as sequential hold & movement segments with features (Liddell 1984)
- Explicit claim: Only difference between sign and speech is size & content of feature system
- Rawski 2017: metathesis, compound reduction, partial reduplication are ISL in speech/sign

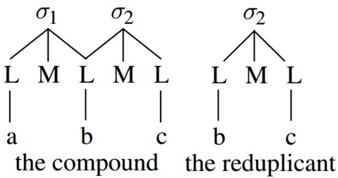
ASL Final Syllable Reduplication



pics from Sandler & Lillo-Martin 2006



ASL Final Syllable Reduplication

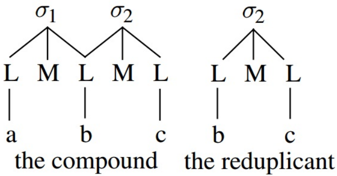


4-ISL function: ∅ ! LML / LML__n

Input String: LMLML

o L M L M L n

ASL Final Syllable Reduplication

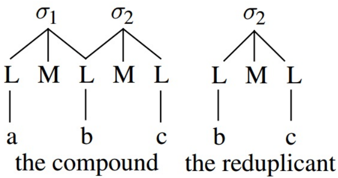


4-ISL function: ∅ ! LML / LML__n

Input String: LMLML

o L M L M L n

ASL Final Syllable Reduplication



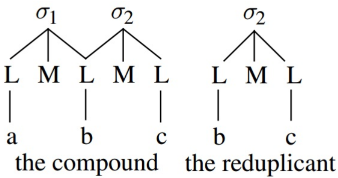
4-ISL function: ∅ ! LML / LML__n

Input String: LMLML

o L M L M L n

L

ASL Final Syllable Reduplication



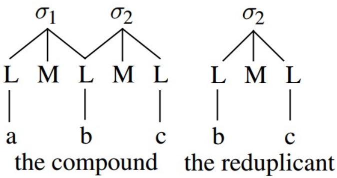
4-ISL function: ∅ ! LML / LML__n

Input String: LMLML

o L M L M L n

L

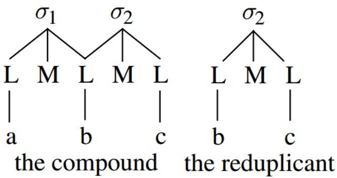
ASL Final Syllable Reduplication



4-ISL function: $\emptyset$! LML / LML__n
Input String: LMLML

o L M L M L n
L M

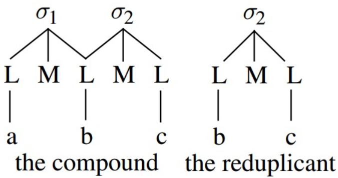
ASL Final Syllable Reduplication



4-ISL function: ∅ ! LML / LML__n
Input String: LMLML

o L M L M L n
L M

ASL Final Syllable Reduplication

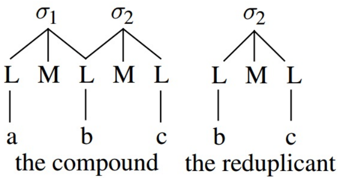


4-ISL function: ∅ ! LML / LML__n

Input String: LMLML

o	L	M	L	M	L	n
	L	M	L	M		

ASL Final Syllable Reduplication



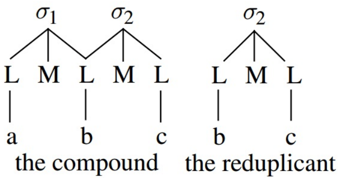
4-ISL function: $\emptyset$! LML / LML__n

Input String: LMLML

o L M L M L n

L M L M

ASL Final Syllable Reduplication

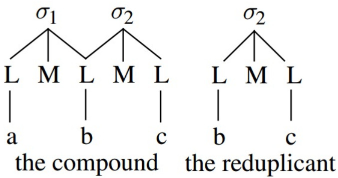


4-ISL function: $\emptyset$! LML / LML__n

Input String: LMLML

o	L	M	L	M	L	n
	L	M	L	M		

ASL Final Syllable Reduplication



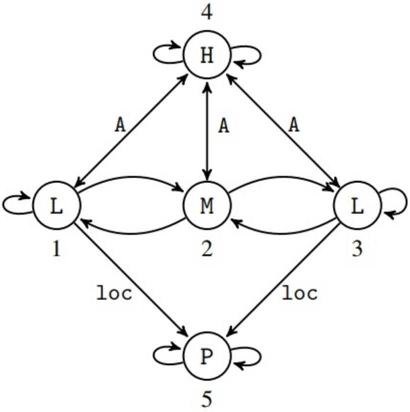
4-ISL function: $\emptyset$! LML / LML__n

Input String: LMLML

o	L	M	L	M	L	n
	L	M	L	M	L	LML

Signs as Graphs

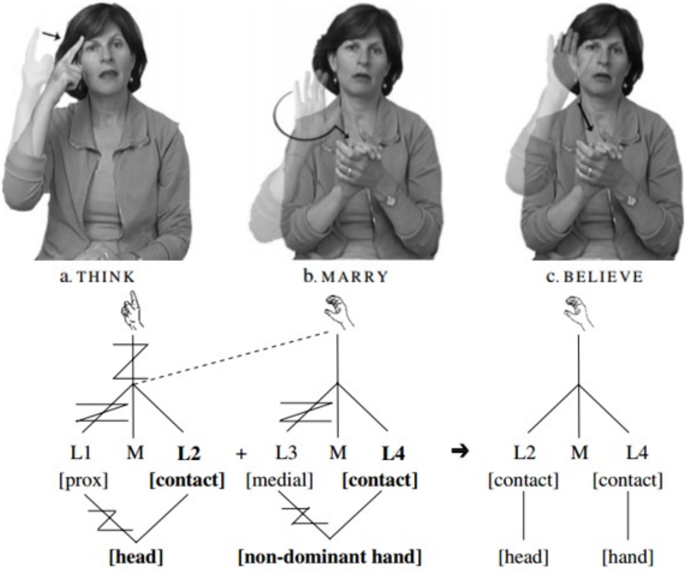
IDEA (ASL)



- encodes autosegmental relations (Sandler 1989, van der Hulst 1993, Brentari 1998)

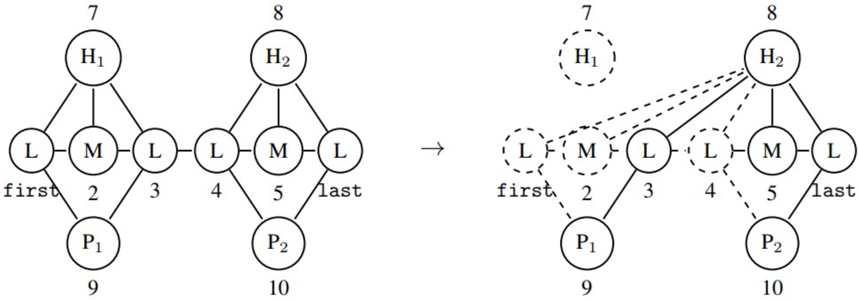
picture from (Sandler & Lillo-Martin 2006)

Compound Reduction



pictures from (Sandler & Lillo-Martin 2006)

Compound Reduction



pictures from (Sandler & Lillo-Martin 2006)

ISL Across Speech and Sign

Process	Strings	Graphs
Metathesis	ISL	A-ISL
Partial Reduplication	ISL	A-ISL
Compound Reduction	ISL	A-ISL

Interpretation

- Strict Locality is salient across spoken and signed phonology
- Locality ranges over the representation
- "Adapted systems" view
  - signers exploit nonlinear structure
  - restrict sequential structure to preserve ISL requirements
- Locality as unified inductive learning bias

Learning Input Strictly Local Functions

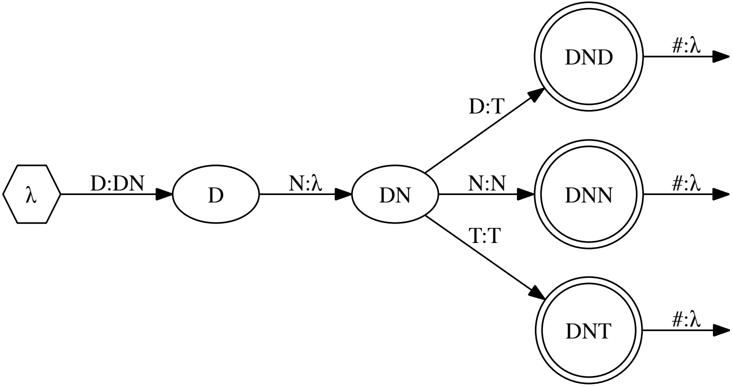
ISL Function Learning Algorithm (Chandlee et al 2014)

- Input strict locality as an inductive principle
- Generalize ISL functions from positive data.
- Successful Identification in the Limit (Gold 1967)

State merging approach

Final devoicing: f(DND, DNT), (DNN, DNN), (DNT, DNT)...g

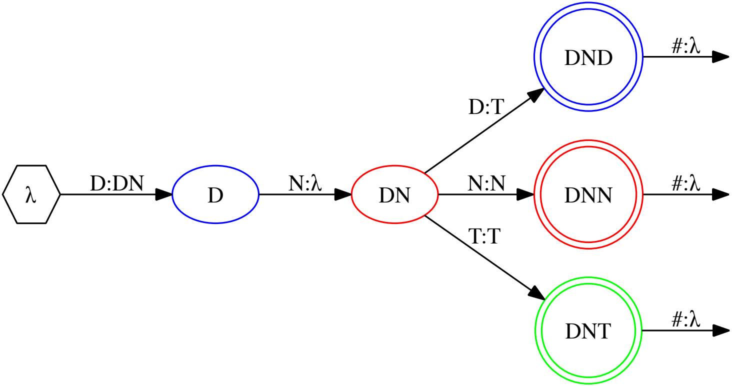
1. Prefix Tree Representation of data
2. Merge states that have the same suffix of length k 1.



State merging approach

Final devoicing: f(DND, DNT), (DNN, DNN), (DNT, DNT)...g

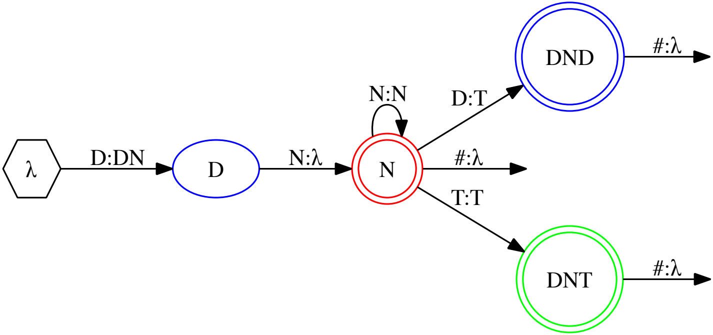
1. Prefix Tree Representation of data
2. Merge states that have the same suffix of length k 1.



State merging approach

Final devoicing: *f*(DND, DNT), (DNN, DNN), (DNT, DNT)...*g*

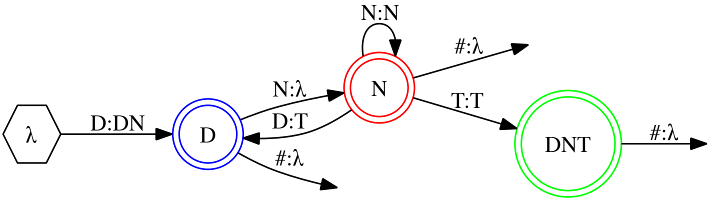
1. Prefix Tree Representation of data
2. Merge states that have the same suffix of length *k*   1.



State merging approach

Final devoicing: f(DND, DNT), (DNN, DNN), (DNT, DNT)...g

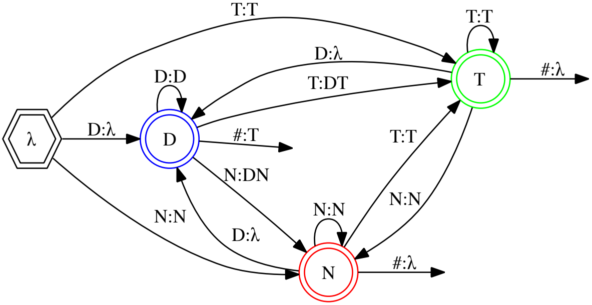
1. Prefix Tree Representation of data
2. Merge states that have the same suffix of length k 1.



State merging approach

Final devoicing: f(DND, DNT), (DNN, DNN), (DNT, DNT)...g

1. Prefix Tree Representation of data
2. Merge states that have the same suffix of length k 1.



Cognitive Implications

Our proposal: Learners are amodally sensitive to

- particular locality representations (e.g. substrings/subgraphs)
- ISL memory restrictions (e.g. bounded substrings).
  - reflected experimentally (Finley 2011; Lai 2015, Avcu 2017)

Any cognitive mechanism with ISL complexity is sensitive to length k blocks of consecutive events occuring in the underlying structure. (Rogers et al 2013)

If structures occur in time, this means sensitivity, at each point, to immediately prior sequence of $k - 1$ events.

All learning systems necessarily structured by representational and computational nature of their domains (Rawski & Heinz 2019).

Answering Poeppel's "Mapping Problem"

Maps problem: Find brain areas correlating with cognitive tasks
Mapping problem: Decompose cognition into neuronal operations

Poeppel 2012

"focus on the operations and algorithms that underpin language processing"

"commitment to an algorithm or computation in this domain commits one to representations of one form or another with increasing specificity and also provides clear constraints for what the neural circuitry must accomplish."

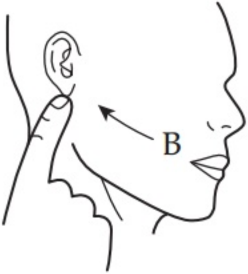
Where to go from here?

- Further linguistic work on computational comparisons of phonology across speech and sign
- Learning algorithms/theorems, integrating learnability with linguistic theory
- Representational tradeoffs (strings, trees, graphs, etc)
  - Phonology, morphology, syntax, may be local over the right representations (Graf et al 2018).
- Computationally motivated experimental work on modality

Metathesis

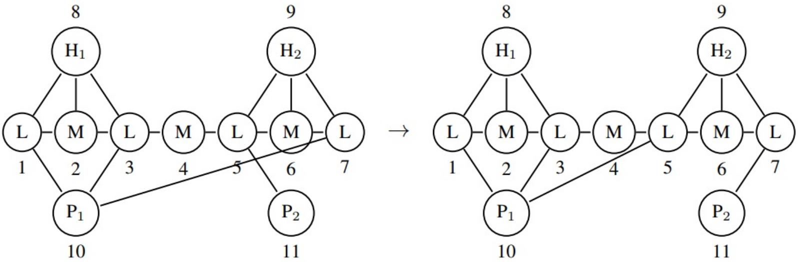
a. FATHER DEAF (ASL)

b. MOTHER DEAF (ASL)



pictures from (Sandler & Lillo-Martin 2006) and Wilbur (2012)

Metathesis



pictures from (Sandler & Lillo-Martin 2006) and Wilbur (2012)