SEARCHING FOR MORPHOLOGICAL PRODUCTIVITY

Sarah Payne[1], Caleb Belth[2], Jordan Kodner[3], Charles Yang[1]

[1]University of Pennsylvania
[2]University of Michigan
[3]Stony Brook University

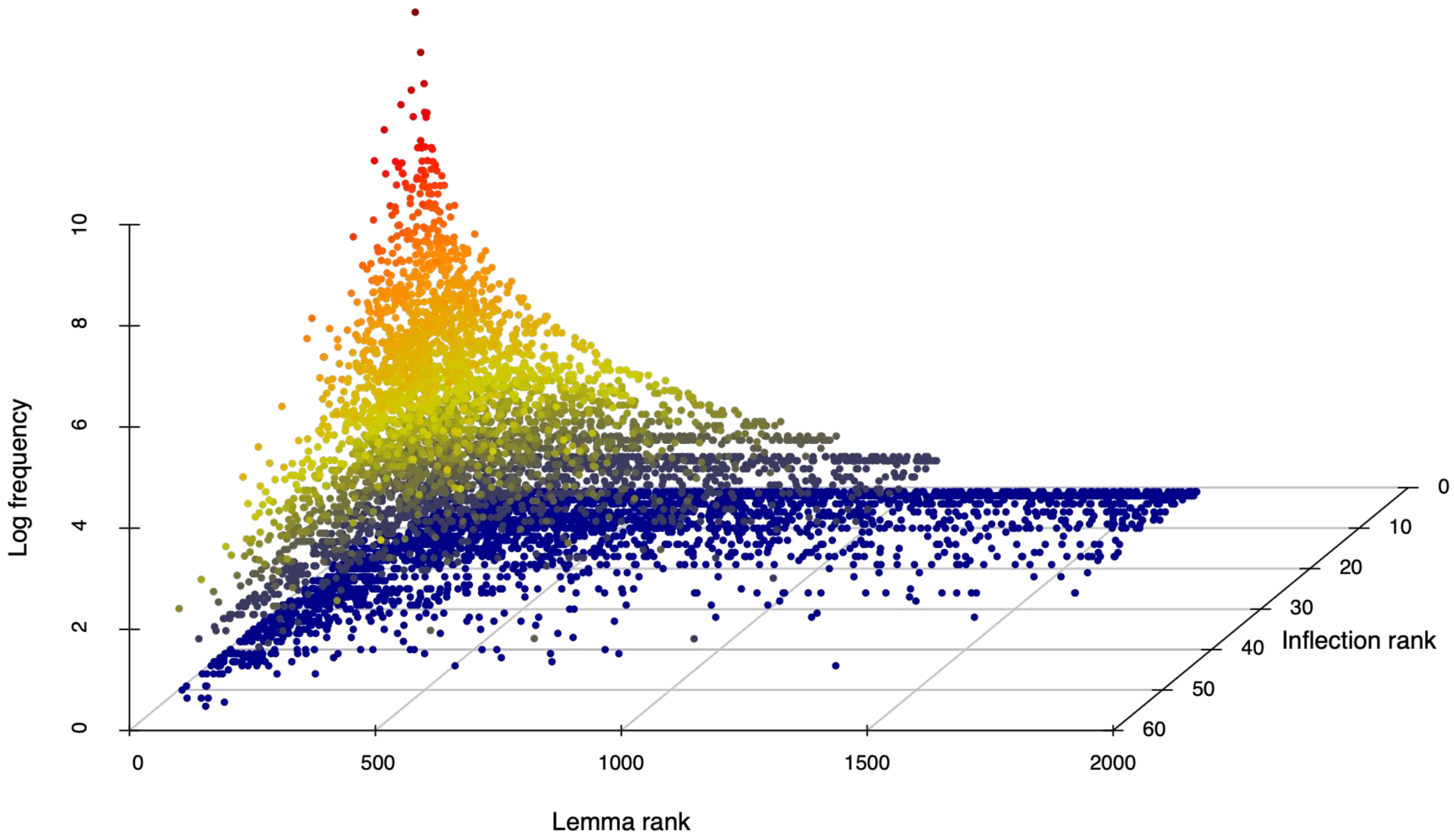
LSA 2022 Annual Meeting

SEARCHING FOR PRODUCTIVITY

How do children discover productive generalizations?

- Overcoming sparsity
- Despite exceptions
- When multilayered

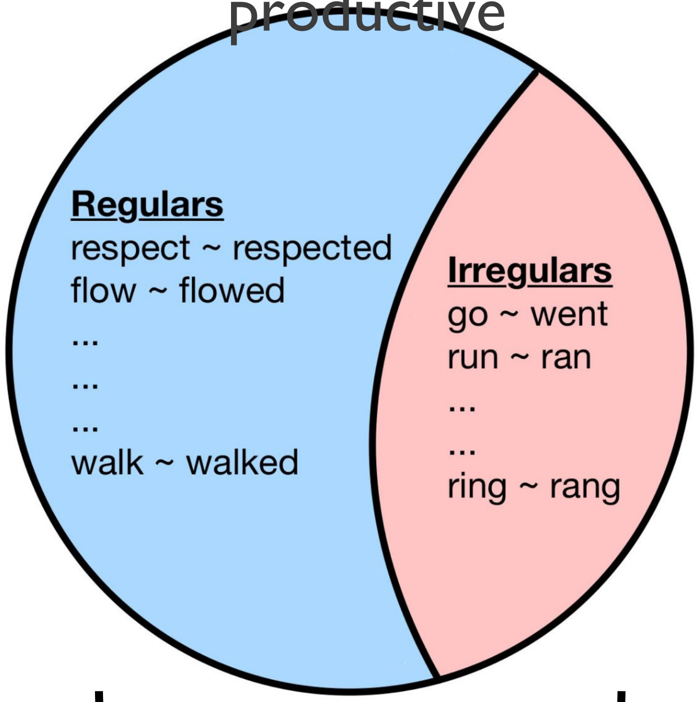
BACKGROUND: SPARSITY



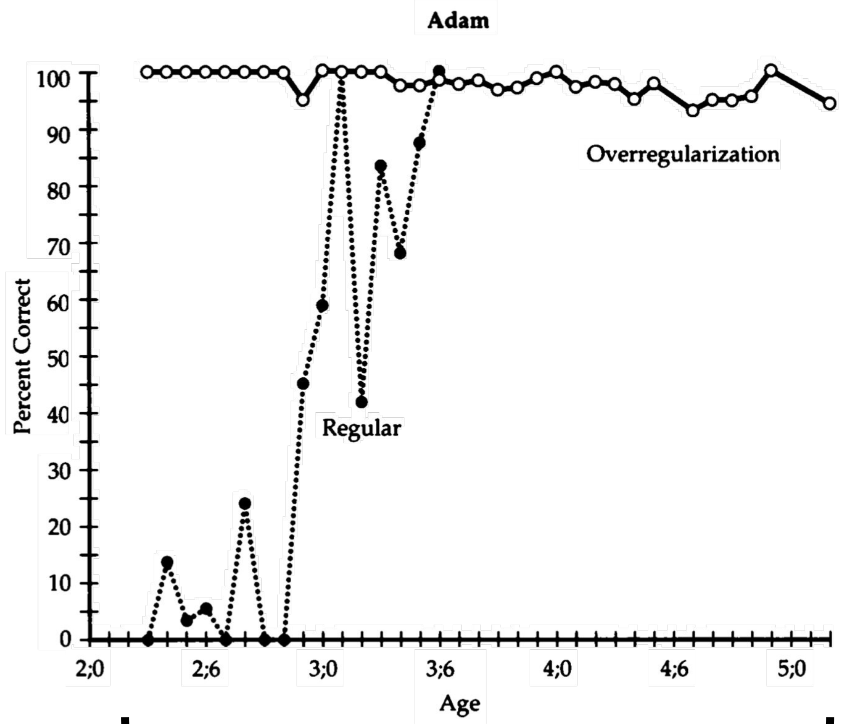
Courtesy of Erwin Chan & Constantine Lignos

BACKGROUND: PRODUCTIVITY

English Past Tense: Statistically dominant rule = productive



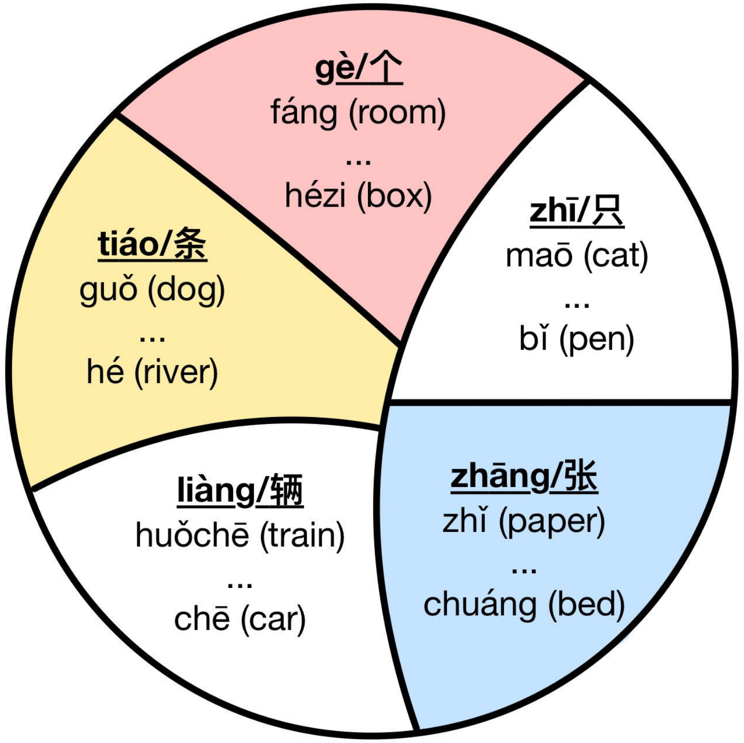
Productivity Despite Exceptions



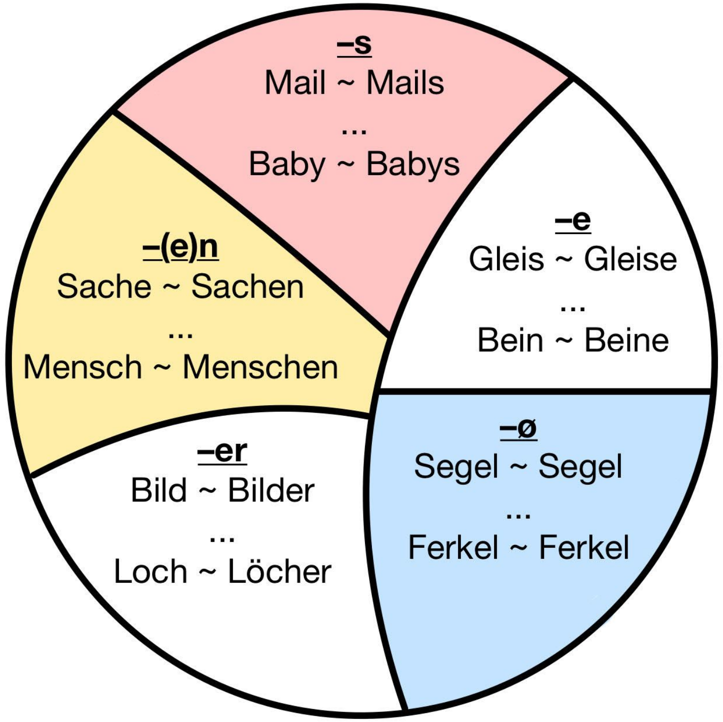
"U-shaped" learning

BACKGROUND: PRODUCTIVITY

German Plurals & Mandarin Classifiers: Restricted to subgroups



Semantics



Gender & Phonology

CONTRIBUTIONS

We present a **model of morphological learning** capable of extracting **linguistically interpretable rules** from developmentally plausible vocabularies

DATA

Input: **(lemma, inflected, feature)**

English: (walk, walked, {3, SINGULAR, PAST})
German: (Sache, Sachen, {FEMININE})
Chinese: (rén, gè rén, {+ANIM, +CONC, -FLAT, +HUM, +NAT, -SLEN, -VEH})

	English Past Tense	German Plurals	Mandarin Classifiers
Max Training Size	600 words	360 words	100 words

MODEL: THE TOLERANCE PRINCIPLE

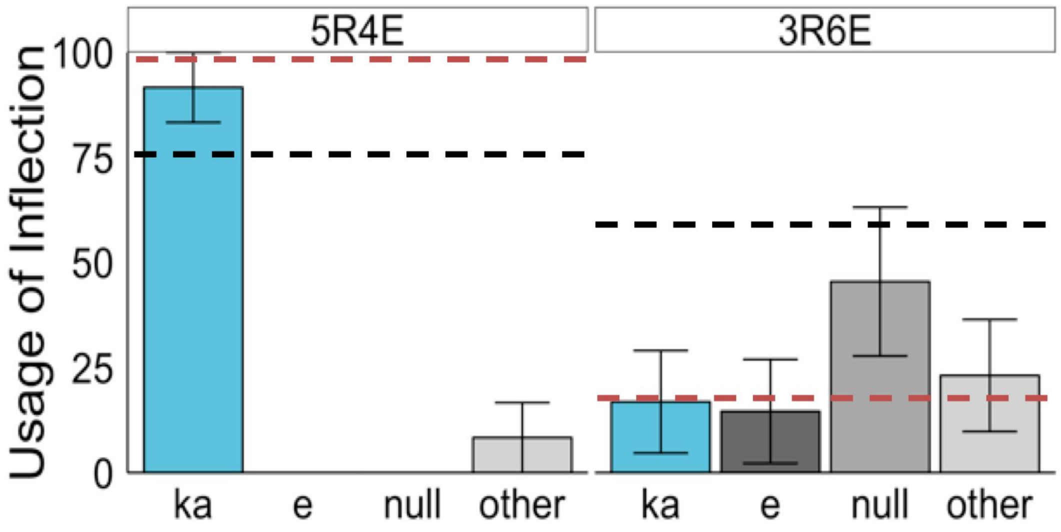
- **Intuition:** given a set of items:
  - If **many** do X, then all do X (generalization)
  - If **few** do X, then remember the few that do (lexicalization)
- Threshold defined by efficiency:

$$\underbrace{e}_{\text{exceptions}} \leq \theta_N = \underbrace{\frac{N}{\ln N}}_{\text{threshold}}$$

MODEL: THE TOLERANCE PRINCIPLE

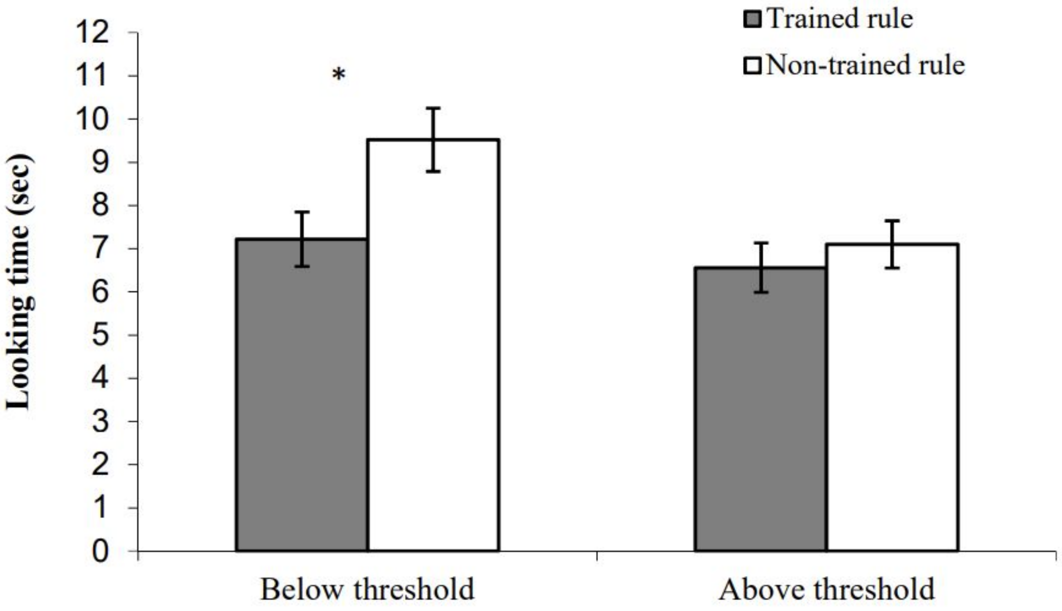
Empirical evidence from artificial language studies

15 children age 6-8 years



Schuler, Yang & Newport (2016, Submitted)

$$\theta_9 = 4.2$$



Emond & Shi (2021)

$$\theta_{16} = 5.7$$

MODEL: ABDUCTIVE SEARCH

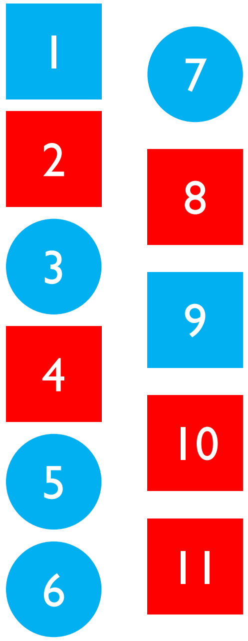
TP applied **recursively:**

- Try forming **rule** over **set** of **N** items
- If **rule** not productive, subdivide **set** into disjoint subsets
- Repeat within each subset

Terminates when:

- Productive rule found (generalization)
- Or, no more subdivisions possible (lexicalization)

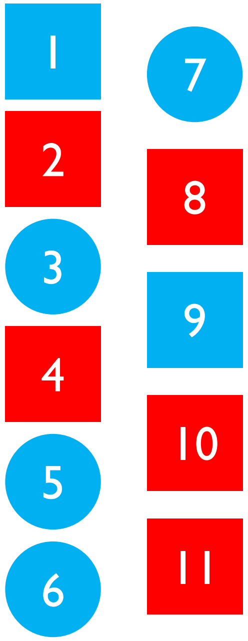
MODEL: ABDUCTIVE SEARCH



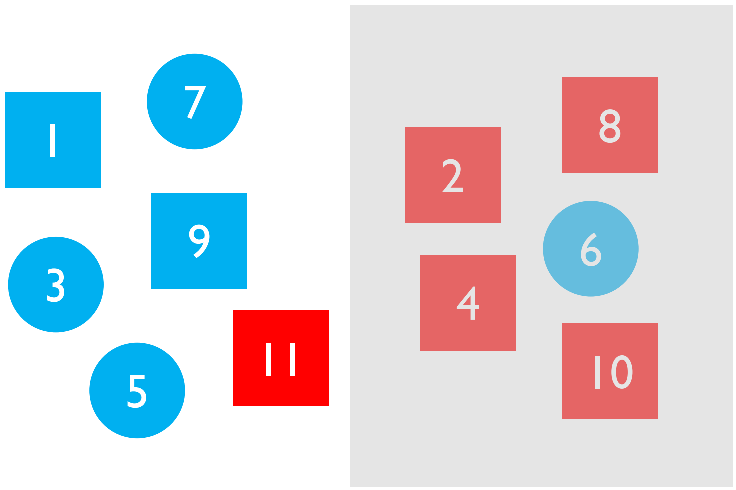
- Find the most frequent color (6 vs. 5)
- **Hypothesize** a rule {Features} →
  - Odd→
- **Test** the rule "Odd→ "
  - *TP check (N=6, e=1)*: 1, 3, 5, 7, 9, 11
- R1 productive: Odd→ Exceptions 11
- Recurse over remaining items
- R2 productive: Even → Exceptions 6

MODEL: SELECTING A FEATURE

Multiple ways to subdivide **N** items:

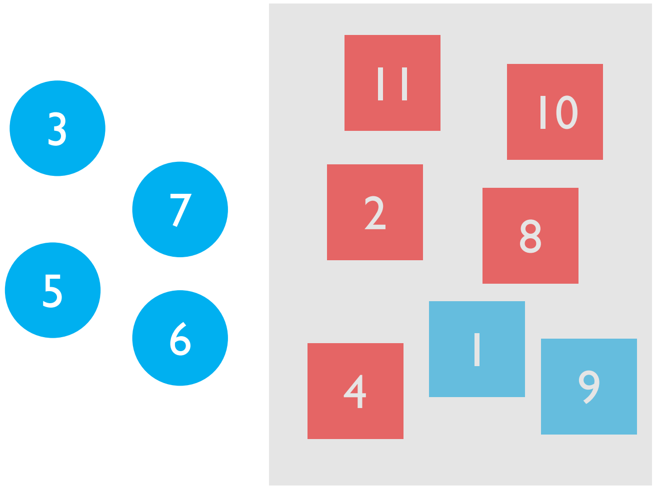


11 Items



<u>Most **Frequent** [1]</u>

Odd → Blue



<u>Most **Consistent** [2,3]</u>

◯→ Blue

[1] Carla L Hudson Kam and Elissa L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1(2):151–195.
[2] LouAnn Gerken. 2006. Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*, 98(3):B67–B74.
[3] Patricia A Reeder, Elissa L Newport, and Richard N Aslin. 2013. From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive psychology*, 66(1):30–54.

RESULTS

Q1: How accurately does our model learn morphology?

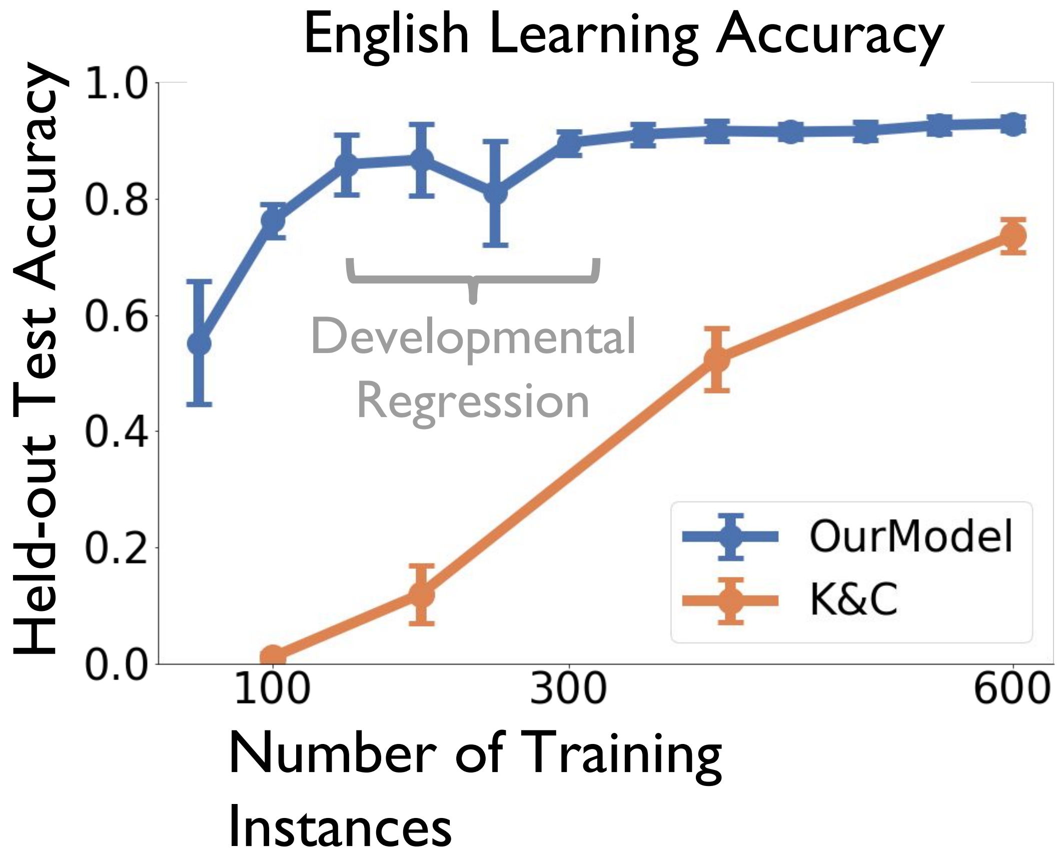
- English past tense
- German plurals
- **Comparison:** Kirov and Cotterell (2018)'s Neural Network model (K&C)

Q2: Are the results developmentally plausible?

- English past tense learning trajectory
- Linguistic interpretability of rules
- Attends to relevant features

RESULTS: ENGLISH PAST TENSE

Child-like developmental trajectory

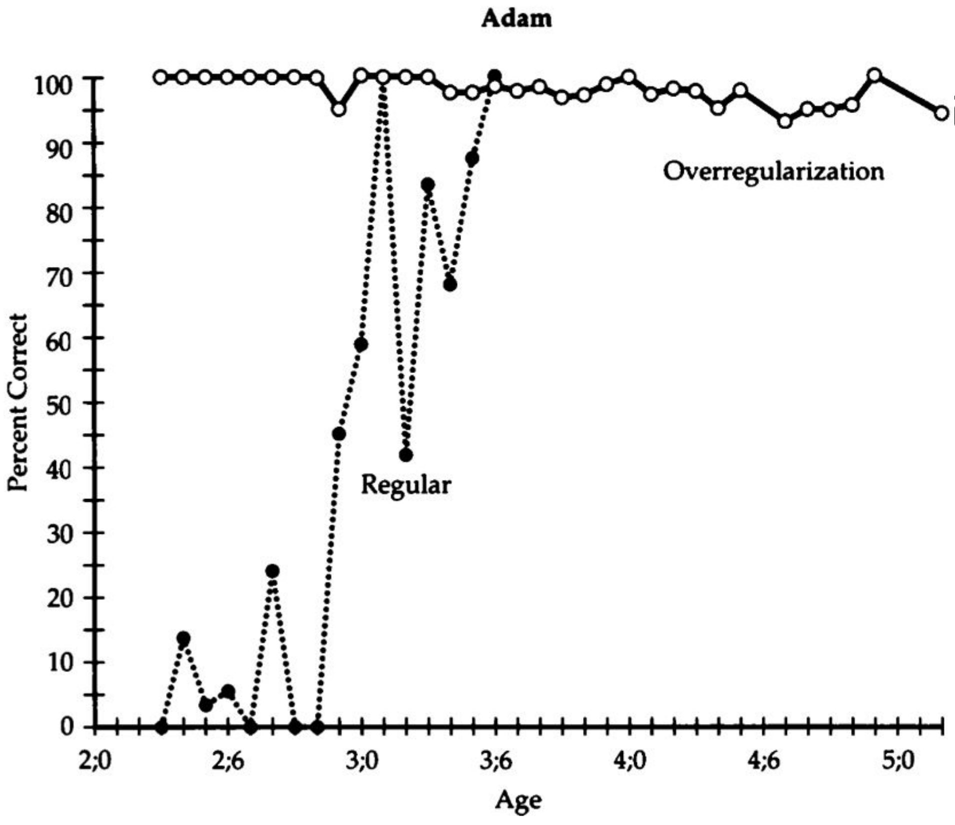


Linguistically interpretable rules

[+VE

[+VE

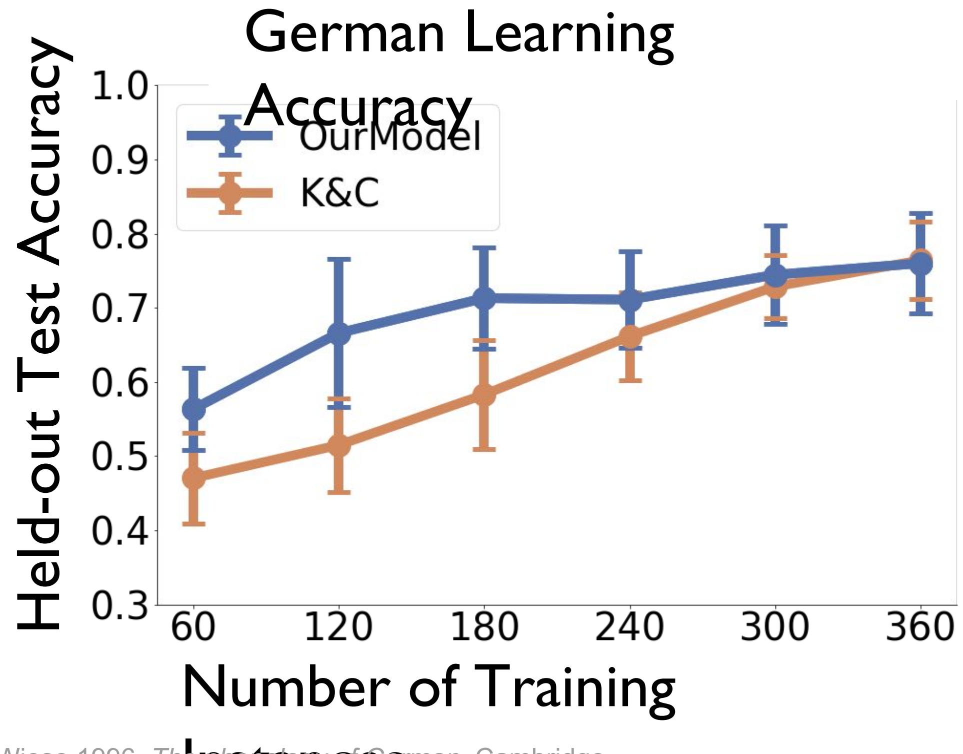
[+VE



#

RESULTS: GERMAN PLURALS

Child-like developmental trajectory



Linguistically interpretable rules

Gender | Phonology

[M] → -e / [d, b, f, g, h, z, s, k, m, …]#
[F] → -en / [r, t, n, g, z, hl]#
[F] → -n
…
[N] → -e / [r, l, m, t, n, z]#
…

-(e)n

[4] Wiese 1996. *The phonology of German.* Cambridge.
[5] McCurdy, K., Goldwater, S., & Lopez, A. (2020). Inflecting when there's no majority: Limitations of encoder-decoder neural networks as cognitive models for german plurals. In Proceedings of the 58th annual meeting of the association for computational linguistics, ACL (pp. 1745–1756).

RESULTS: MANDARIN CLASSIFIERS

[-Veh, -Slen, -Flat, -Hum, -Anim, +Conc, -Nat] → gè (個/个)

[-Veh, -Slen, -Flat, -Hum, -Anim, +Conc, +Nat] → gè (個/个)

[-Veh, -Slen, -Flat, -Hum, -Anim, -Conc, -Nat] →gè (個/个)

…

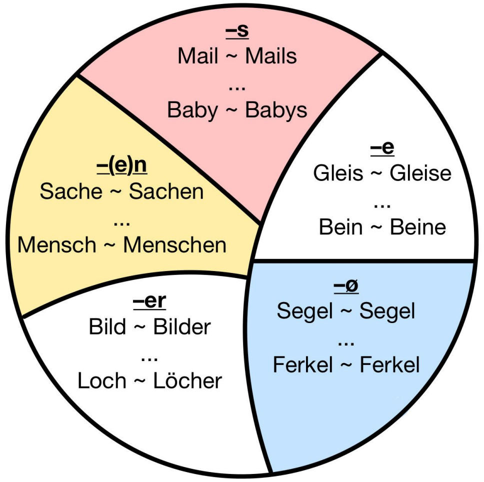
[-Anim, -Hum, -Nat, +Conc, -Slen, -Flat, +Veh] → liàng (輛/辆)

- **Semantic conditions learned** ✓
- **Irrelevant phonological properties ignored [5]** ✓

[5] Hu 1993. The acquisition of Chinese classifiers by young Mandarin-speaking children. BU dissertation.

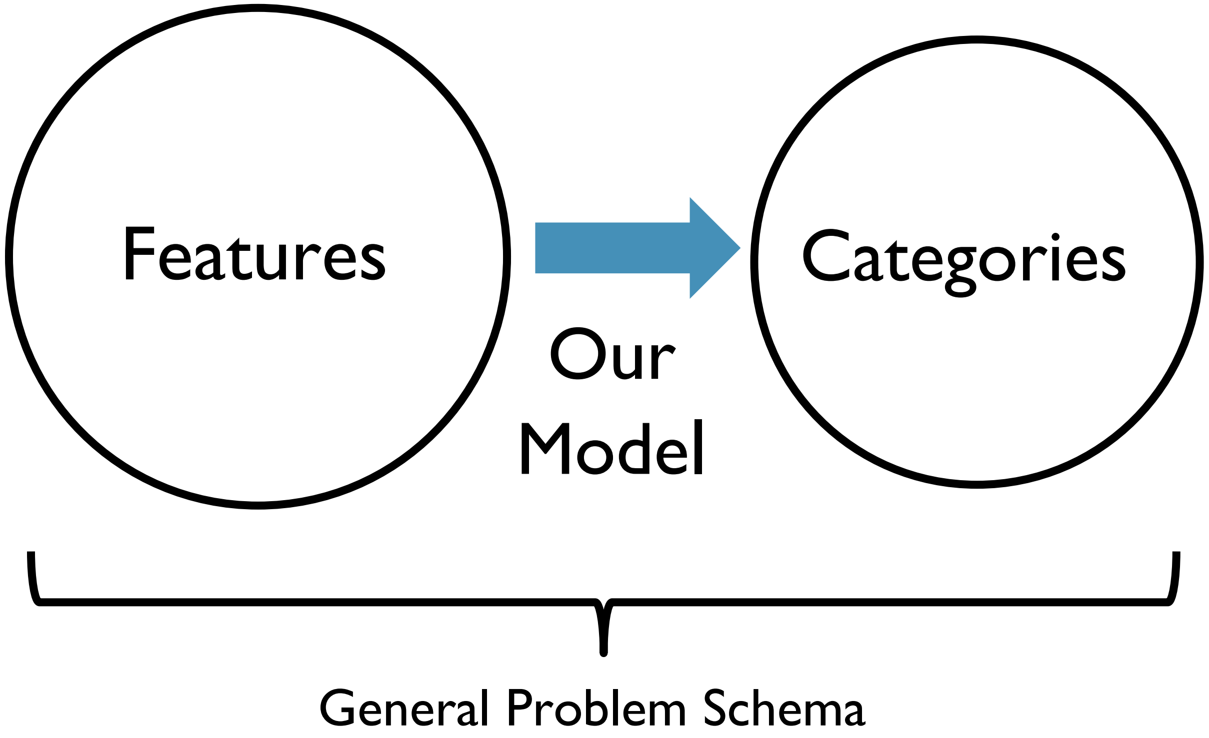
CONCLUSION

- Abductive, recursive search + TP provides plausible account of morphological acquisition
- Lexicon partitioned into categories
  - The rules yielding these categories are a ‘good enough’ grammar
  - That is, they regularize learning
- Preserves explicit distinction between
  - generalization and lexicalization [6, 7]
  - walk →walked vs. run →ran



[6] Berko, J. (1958). The child’s learning of english morphology. Word, 14(2-3), 150–177.
[7] Lignos, C., & Yang, C. (2016). Morphology and language acquisition. In G. Hippisley Andrew R. abd Stump (Ed.), The Cambridge handbook of Morphology (p. 765-791). Cambridge: Cambridge University Press.

CONCLUSION



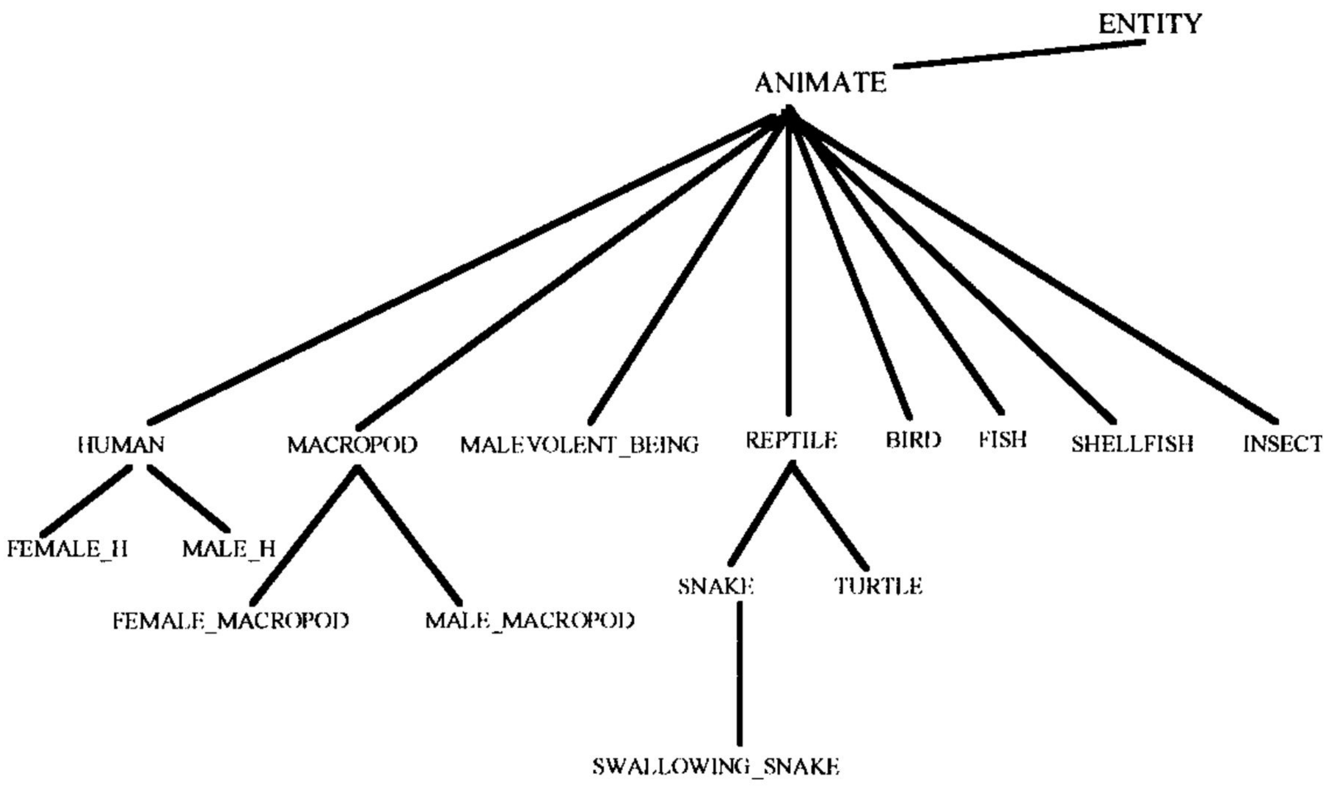
Applicable to linguistic mappings beyond morphology

e.g.,

- {Phonology, Semantics} → Gender
- {Distributional Properties} → Phonological Natural Classes

…

CONCLUSION

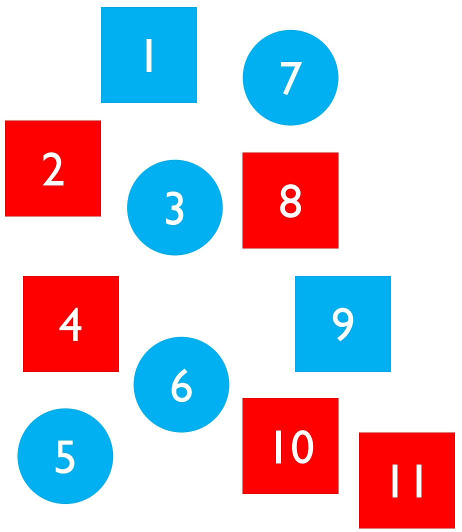


Animate Semantic Hierarchy Mayali

Evans, Brown & Corbett (2002/2019:132)

Maybe general category formation process?

- Object Properties→ Categories



THANK YOU!!!

Thanks to Deniz Beser, John Trueswell and his lab, and the members of LING-570 and LING-300 at the University of Pennsylvania