The Greedy and Recursive Search for Morphological Productivity

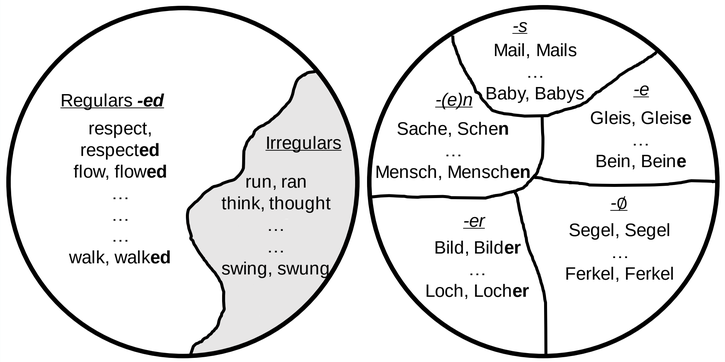
Caleb Belth[1], Sarah Payne[2], Deniz Beser[3], Jordan Kodner[4], Charles Yang[2]
[1]University of Michigan, [2]University of Pennsylvania, [3]Information Sciences Institute, USC, [4]Stony Brook University

Background

- Language acquisition involves discovering productive, rule-like knowledge

Productivity

- Productivity arises in the face of exceptions (a)
- Productivity may be restricted to subgroups (b)



(a) English Past Tense (b) German Noun Plurals

Sparse Input

- Children discover main generalizations by the time their vocabulary is a few hundred words

Contributions

- Cognitively motivated model, ATP
- Developmentally realistic experiments

Prior Work

	Tolerance Principle (TP)	Neural Models
Predicts productivity	✓	✗
On Naturalistic Data	✓	✗
Automatically generates hypotheses	✗	✓

Selected References
1. 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).
2. Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. MIT press.
3. Wiese, R. (1996). *The phonology of German*. Oxford: Clarendon.
4. Xu, F., & Pinker, S. (1995). Weird past tense forms. *Journal of Child Language*, 22(3), 531-556

Our model, ATP, **automatically** identifies the **productive rules** in a language's morphology. It **abductively** refines its hypothesized knowledge and **mirrors** many empirical facts from the acquisition literature.

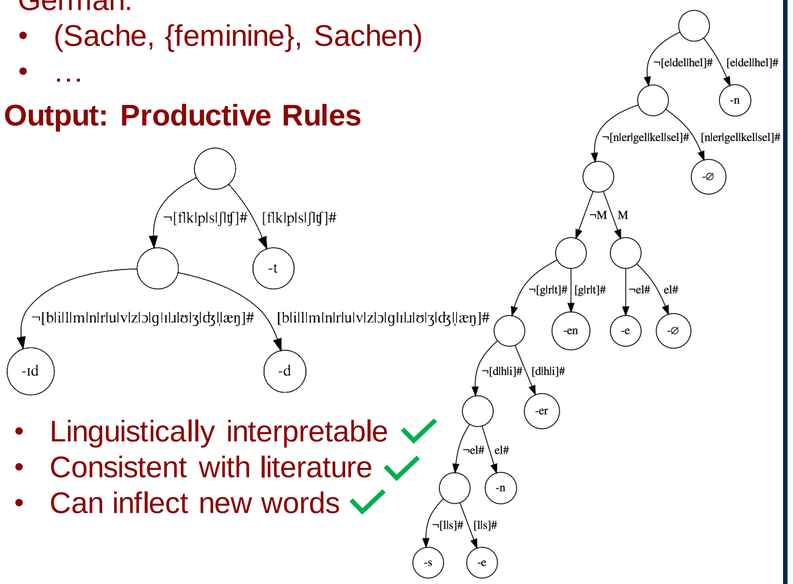
ATP: Proposed Model

Input: (lemma, features, inflection) pairs

English:
- (walk, {3rd person, singular, past tense}, walked)
- …

German:
- (Sache, {feminine}, Sachen)
- …

Output: Productive Rules

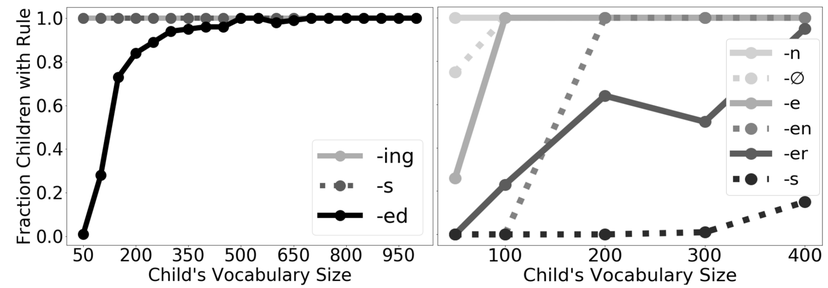


- Linguistically interpretable ✓
- Consistent with literature ✓
- Can inflect new words ✓

Procedure: Recursively Apply the Tolerance Principle
- Hypothesize rule and check vs. the TP: $e \leq \frac{N}{\ln N}$
- If it fails, subdivide based on a feature and recurse

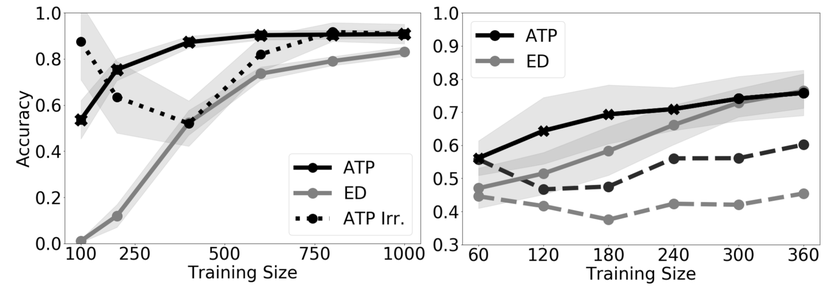
Results

ATP's **Order of Acquisition** closely follows child development



(a) English (b) German

ATP's **Accuracy** surpasses a modern neural network



(a) English (b) German

ATP's **Nonce Word** productions correlate highly with human's

	%R	%NR	ρ
-(e)n	0.19	0.23	**0.43**
-e	0.45	0.62	0.01
-∅	0.07	0.00	**0.55**
-er	0.29	0.0	**0.46**
-s	0.01	0.15	**0.64**
other	0.00	0.00	

Conclusions and Future Work

- ATP can be extended to linguistic generalizations beyond morphology
- ATP's code can be run on your own data: https://bit.ly/3cTaqOd