Theory and predictions for the development of morphology and syntax: A Universal Grammar + statistics approach

LSCI 151C/251C

Moldir Baidildinova

Outline

- What is UG + stats?
- Why UG is part of developmental theories
- (The statistics component)
- (Examples from morphosyntax development)
- Summary and conclusion

What is UG + stats?

- Combines **innate language-specific knowledge** (UG) with **statistical learning**
- Implemented through **computational cognitive modeling**
    - **UG part:** *The linguistic representations the child is learning (e.g., syntactic structure, parameter settings, etc.)*
    - **Stats part:** *the acquisition process the child undergoes*
    - **'proof of concept'**

Why Use Computational Models?

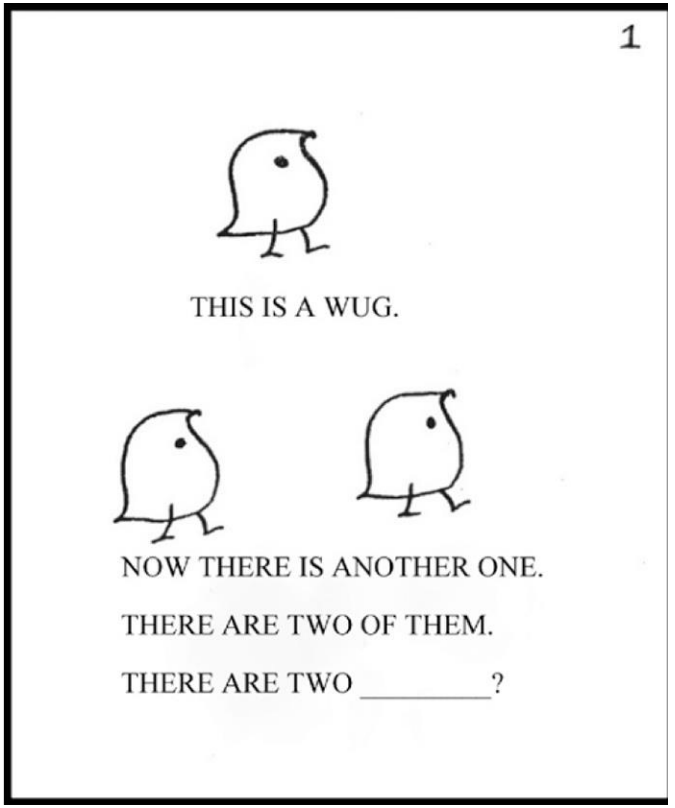
- Concrete implementation of theories
- Allows generation of precise, **testable** predictions
- Evaluated against empirical **child behavior data**

Model implementation involves:

1. Embedding the relevant prior knowledge
2. Learning mechanisms (statistical learning)
3. Realistic Input (child-directed speech)
4. Generating Output (predicted behavior)

UG Component (Section 2.1)

- Motivation: Solve **the Poverty of the Stimulus**
  - *the input data often appears too limited to explain how efficiently children identify the correct linguistic rules.*
  - *Berko, J. (1958). The Child's Learning of English Morphology. Word, 14, 150-177*.



UG Component (Section 2.1)

- UG = innate + language-specific knowledge
- UG structures child's hypothesis space:
  - Defines linguistic building blocks (e.g., parameters)
  - Word order:
    - English: *Anne ate apples.*
    - Kazakh: *Anna alma zhedi.*
- Internal bias helps reduce the hypothesis space, thus enabling children to succeed at language acquisition.

Statistics Component (Section 2.2)

- Statistical learning = counting & inference
- Learner must know what to count → defined by UG
- Inference mechanisms:
- Bayesian inference
- Reinforcement learning
- Tolerance & Sufficiency Principles (Yang 2005, 2016)

Bayesian Inference (2.2.1)

- Learner chooses the most probable hypothesis given data D:
- Posterior = Prior * Likelihood / Evidence
- Learners update beliefs as they receive more input
- Supported by child data (as young as 6 months)

Reinforcement Learning (2.2.2)

- Based on reward/penalty for hypothesis options
- Common implementation: linear reward-penalty scheme
- Supported by child operant learning data (from 10 weeks old)

Tolerance & Sufficiency (2.2.3)

- Child adopts rule only if exceptions are tolerable (Tolerance Principle)
- Must be enough rule-following examples (Sufficiency Principle)
- Applied to morphology, e.g., English past tense
- Supported by evidence in 5- to 8-year-olds (Schuler et al. 2016)

Conclusion

- UG+stats models provide a formal framework for testing developmental theories
- Combining UG constraints with statistical mechanisms yields testable, interpretable predictions
- Promising results in modeling morphological and syntactic development
- Future work: Increase precision of predictions (e.g., age of acquisition)

Citation

- Pearl, L. (2020). Theory and predictions for the development of morphology and syntax: A Universal Grammar + statistics approach. Ms., University of California, Irvine.