How do Learners Identify Movement Dependencies?

Learners need to identify both local and non-local syntactic dependencies:

1. Recognize local argument relations, sensitive to verb transitivity [1, 2, 3]
2. Use verb knowledge to cheat at interpreting wh-dependencies [4, 5]
3. Reliably produce and comprehend wh-questions [4, 6, 7]

Empirical Background

W/Dependency Development in Infancy

15- to 16-month-olds: Recognize local argument relations, sensitive to verb transitivity [1, 2, 3]
May use verb knowledge to cheat at interpreting wh-dependencies [4, 5]

18- to 20-month-olds: Know that a moved wh-phrase is an argument in a wh-question [3]
Reliably produce and comprehend wh-questions [4, 6, 7]

Infants represent argument movement after learning verb argument structure

Hypothesis: Gap-Driven Learning

Learners may use verb argument structure knowledge to identify different types of movement dependencies [3, 4, 5, 7]:
- Notice when expected argument of a verb is missing in its canonical position (gap)
- Identify what forms co-occur with unexpectedly missing arguments
- Infer what underlying dependencies are responsible for these forms
- Current question: Is this hypothesis computationally feasible, given the data that children have to learn from?

Prior Model: Acquiring Argument Structure

It is computationally possible for learners to identify verb argument structure even before they can recognize moved arguments [8, 9]

- Input filtering: assume data has both signal and noise, and learn to filter noise

Syntactically-Informed Distributional Analysis

Learners might combine verb argument structure knowledge with distributional learning to identify which forms characterize movement in English

Joint Inference:
- Categorize sentences according to their surface forms
- Use verb transitivity knowledge to infer which sentence ‘categories’ contain object gaps

Features of Argument-Gap Categories

Distinctive features of argument-gap categories included forms that characterize movement in English, but also included irrelevant forms

Data

18,503 sentences of child-directed speech from the CHILDES Treebank [10]
- 50 frequent transitive, intransitive, and alternating verbs learned by prior model [8, 9]
- Coded for presence of overt direct object and other morphosyntactic features
- For evaluation, also coded for underlying clause type (basic, wh-question, passive, etc.)

Results

Accuracy on Identifying Movement

Model inferred 35 total sentence categories, 15 containing argument gaps

- High overall cluster purity (0.76) compared to actual underlying clause types
- Above-chance accuracy on identifying sentences with movement, higher accuracy on object movement
- Similar accuracy across verb classes

Features of Argument-Gap Categories

Distinctive features of argument-gap categories included forms that characterize movement in English, but also included irrelevant forms

- How do learners ignore spurious correlations in their data?

Discussion

It is possible for a learner to perform distributional learning in order to identify forms that characterize movement dependencies in English

- Provides a computational account for the observed developmental trajectory of argument structure and argument movement acquisition

Distributional learning only goes so far: model identified forms that characterize movement, but also irrelevant forms

- How do learners infer the different dependencies that are responsible?

Current Model

Translational verb
Observations of verb direct object or not?
Category of sentence
Probabilities of direct object
Other features of sentence

Fig. 3 Graphical Model

Mind the Gap: Learning the Surface Forms of Movement

Laurel Perkins, Naomi H. Feldman, and Jeffrey Lidz
University of Maryland, College Park

Acknowledgments

Big thanks to research assistants Liliana Righter, Jordan Schneider, Alexander Shashoua, and John-Paul Tell, as well as Marie Hrazd, Noreet Homa, and Tyler Kaminski. Alexander Williams, Tara Nwaeze, and the Project on Children’s Language Learning. Funding: NSF BCS-1270798, BCS-1501829, OISE-1448615.

References:

Contact: Laurel Perkins, PerkinsL@umd.edu