Bootstrapping into Filler-Gap: An Acquisition Story

Marten van Schijndel
Department of Linguistics
The Ohio State University

March 6, 2012

Overview

Phenomenon

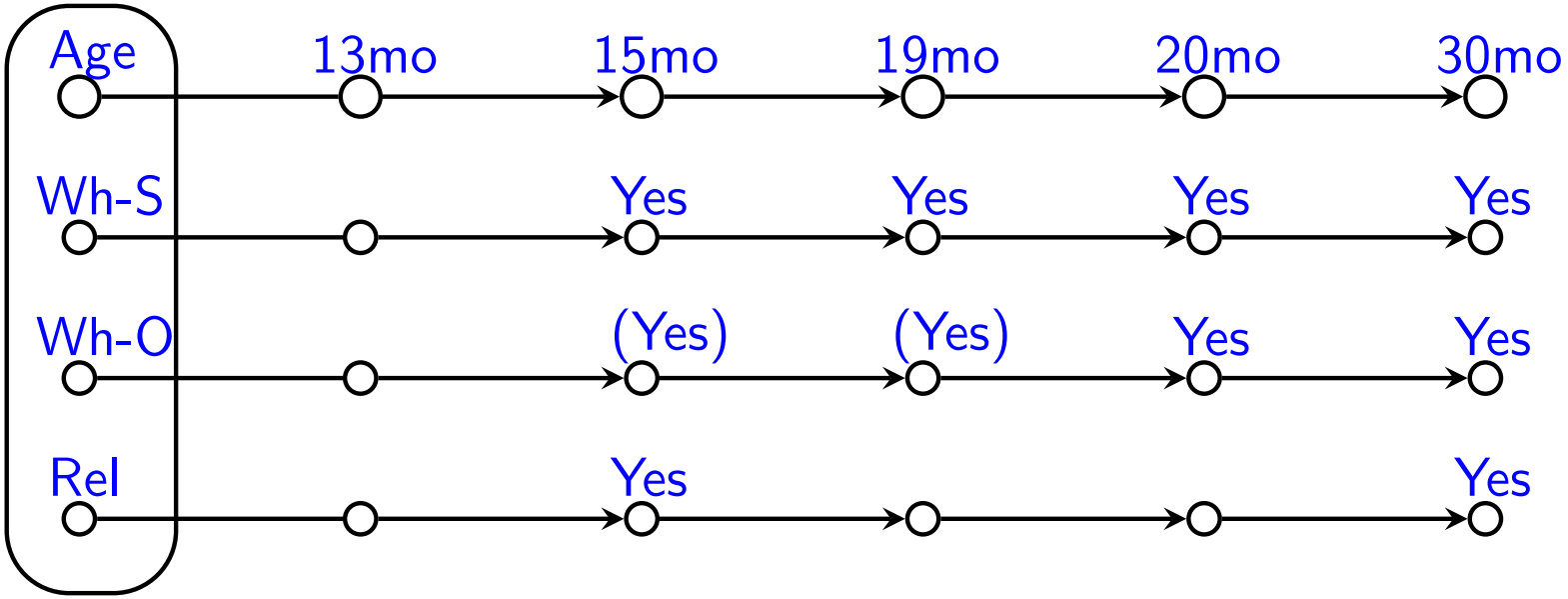
U-shaped curve during filler-gap acquisition

Goal

Model it!

Disclaimer: This talk will largely conflate Subject/Object and Agent/Patient.

Motivation



Developmental timeline of comprehension

Parentheses = marginal comprehension

[Seidl et al., 2003, Gagliardi and Lidz, 2010, Gagliardi et al., 2011]

Rationale

Confusion with dem/det?

- That is a book.
- Gimme that!
- Gimme that book!
- Find the cookie that the mouse ate.

Conflation with questions?

- Who kicked the bucket?
- Who did the burglar assault?
- Find the mouse who the cat ate.

Model

- Structure mapping: nouns used to learn verbs [Yuan et al., 2012]
- Gradual Learning Algorithm [Boersma, 1997]

Assumptions

- Children can identify nouns
  [Shi et al., 1998, Waxman and Booth, 2001]
- 1-to-1 mapping between Ns and roles per sentence
  [Gertner and Fisher, 2012]
- Abstract factors (# N) are used by learners [Xu, 2002]
- Children are sensitive to 'who' and 'that' [Gagliardi et al., 2011]

Implementation Assumptions

- Distributions are Gaussian (fixed-σ)

Inference

Expectation-Maximization

- Initialize with S,V,O = -1,0,1
- Estimate labels using distributions over previous observations
- Estimate new distributions using new data

Evaluation

1. Extract parsed/SRL CDS from Eve corpus [Connor et al., 2008]
2. Collapse A1+ into A1 (object-ness)
3. Run inference
4. Will differing amounts of training data show U-shape?

Results: First Pass

First, let's just see if the data will support the bimodality we'd like.

Results: First Pass

First, let's just see if the data will support the bimodality we'd like.

Role	Before Verb	After Verb
Agent	3826	84
Patient	293	3530

Appearances of each latent variable

Future Work

- Do the project!
- Infer locations of distribution means
- Add ‘that’ and ‘who’ distributions and see where they land
- Add relative clauses and see if differing amounts of training data yield the desired curve

Bibliography I


Boersma, P. (1997).
How we learn variation, optionality, and probability.
Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam, 21:43–58.


Connor, M., Gertner, Y., Fisher, C., and Roth, D. (2008).
Baby srl: Modeling early language acquisition.
In *Proceedings of the Twelfth Conference on Computational Natural Language Learning*.


Gagliardi, A. and Lidz, J. (2010).
Morphosyntactic cues impact filler-gap dependency resolution in 20- and 30-month-olds.
In *Poster session of BUCLD35*.

Bibliography II


Gagliardi, A., Mease, T. M., and Lidz, J. (submitted 2011).
U-shaped development in the acquisition of filler-gap dependencies: Evidence from 15- and 20-month-olds.
Language Learning and Development.


Gertner, Y. and Fisher, C. (2012).
Predicted errors in childrens early sentence comprehension.
Cognition, 124:85–94.


Seidl, A., Hollich, G., and Jusczyk, P. W. (2003).
Early understanding of subject and object wh-questions.
Infancy, 4:423–436.


Shi, R., Werker, J. F., and Morgan, J. L. (1998).
Newborn infants sensitivity to perceptual cues to lexical and grammatical words.
Cognition, 72:B11–B21.

Bibliography III


Waxman, S. R. and Booth, A. (2001).
Seeing pink elephants: Fourteen-month-oldss interpretations of novel nouns and adjectives.
Cognitive Psychology, 43:217–242.


Xu, F. (2002).
The role of language in acquiring object kind concepts in infancy.
Cognition, 85:223–250.


Yuan, S., Fisher, C., and Snedeker, J. (2012).
Counting the nouns: Simple structural cues to verb meaning.
Child Development, 83(4):1382–1399.