Word Learning in the Wild: What Natural Data Can Tell Us

Barend Beekhuizen[1], Afsaneh Fazly[2], Aida Nematzadeh[2], Suzanne Stevenson[2]

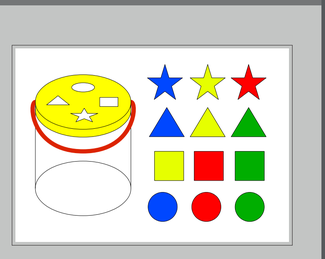
[1] Leiden University, [2] University of Toronto

Word-meaning acquisition

- **Cross-situational word learning**:
  - Learner tracks co-occurrence of words and situationally available meanings
  - Depends on cross-situational availability of meaning
- **Problem**:
  - Computational models of word-meaning acquisition assume a high degree of availability of the cross-situational information for all words
    - Including relational meanings (harder to learn from the situational context than object-labels [2,3])
  - How valid is this assumption?
- **Our approach**:
  - Annotate video data of caregiver-child interaction
  - Investigate situational availability assumption by using computational word learning model

Word learning in the wild: video-taped interaction

- 32 pairs of Dutch mothers and daughters (16mo)
  - playing a game of putting blocks in holes
  - 152 minutes in total
  - 7,500 words of child-directed speech



- For every 3-second interval, annotated features:
  - active game-related objects & participants (child, block)
  - properties of the objects (red, round)
  - the participant's actions on the objects (grab, move)
  - changes in spatial relations among the objects (in, off)
- High inter- and intracoder reliability ($\kappa > 0.8$)

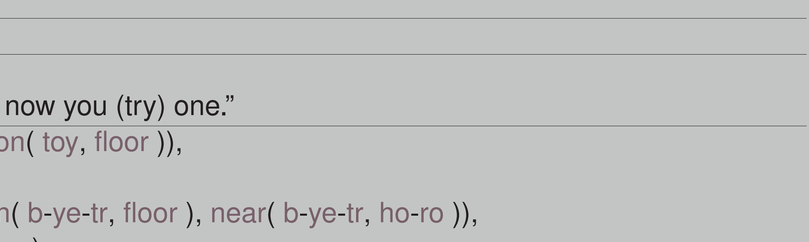
tier	coding/transcription
sit.	<nothing happens>
utt.	*een. nou jij een.*"One. now you (try) one."
sit.	position(mother, toy, on(toy, floor)), grab(child, b-ye-tr), move(child, b-ye-tr, on(b-ye-tr, floor), near(b-ye-tr, ho-ro)), mismatch(b-ye-tr, ho-ro)
utt.	*nee daar*. "No, there."
sit.	point(mother, ho-tr, child) position(child, b-ye-tr, near(b-ye-tr, ho-ro)) mismatch(b-ye-tr, ho-ro)
utt.	*nee lieverd hier past ie niet*. "No sweetie, it won't fit in here."

Table 1: A sample of the dataset. The dash-separated abbreviations denote blocks (b) and holes (ho) and their properties, color (only blocks) ({red,green,blue,yellow}) and shape ({round,star,square,triangular})

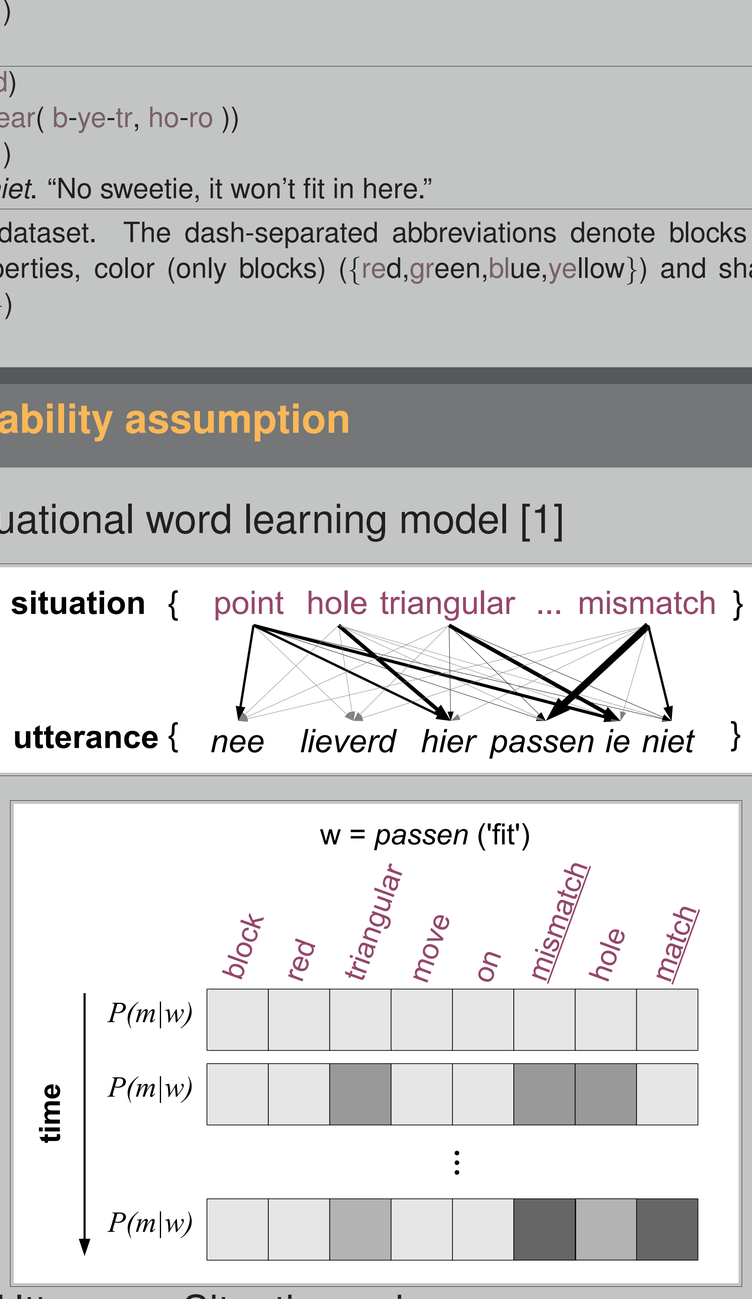
Exploring the availability assumption

- Using a cross-situational word learning model [1]

Aligning features in situation with words



Using alignments to learn probability distributions of features given words



- Trained on 2500 Utterance-Situation pairs

Evaluation

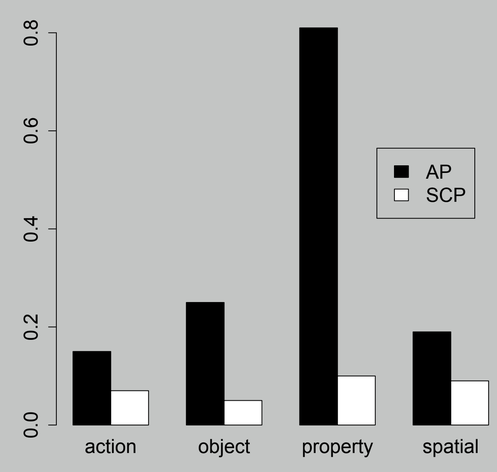
- Evaluated learned probability distributions against hand-annotated relevant features:
  - e.g. *stoppen* - move,in; *blok* - block; *rood* - red; *op* - on
  - four groups of features: action, object, property, spatial
  - evaluation metrics, for each word:

SCP Summed Conditional Probability (probability mass of relevant features)

AP Average Precision (quality of the ranking of relevant features)

Experiment 1

- Situation is the 3 second interval of the utterance



SCP low in general

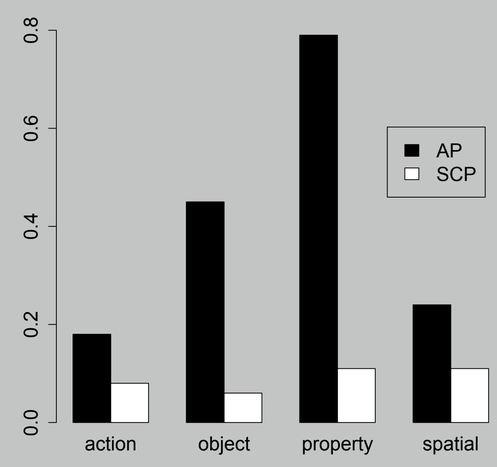
AP poor, except properties

Cross-situational availability is problematic, because of:

- absence of relevant features,
- overwhelming presence of irrelevant features,
- low variability across situations

Experiment 2

- Situation consists of all intervals between the current utterance and the next



SCP remains low

AP general increase

A pragmatically defined attention span

- increases the availability of relevant features for words
- while not increasing the irrelevant features

Key insights

- Developing annotations of naturalistic data is possible
- Cross-situational availability may be low in naturalistic data and the assumption of availability is problematic
- But: results depend on other assumptions (w.r.t. attention and intentions)
- Modeling has to move beyond using mere associations between situations and utterances and look into other mechanisms of word learning in order to understand the mechanisms involved
  - e.g. syntactic bootstrapping, intentions, attention, biases
  - experiment 2: wider attention span increases performance

References

1 A. Fazly, A. Alishahi, and S. Stevenson (2010). A probabilistic computational model of cross-situational word learning. *Cognitive Science*, 34(6):1017–1063.

2 D. Gentner (1978). On Relational Meaning: The Acquisition of Verb Meaning. *Child Development*, 49:988–998.

3 L. Gleitman (1990). Sources of Verb Meanings. *Language Acquisition*, 1(1):3–55.

Acknowledgements: We gratefully acknowledge the funding of BB through NWO of the Netherlands (grant 322.70.001) and AF, AN and SS through NSERC of Canada, and the Faculty of Arts & Science, University of Toronto.

http://dnrb.github.io barendbeekhuizen@gmail.com