Distributional Semantic Models for Vocabulary Acquisition

Raquel G. Alhama, Caroline Rowland, Evan Kidd

TILBURG ◆ UNIVERSITY

MAX PLANCK
M A X
P L A
N C K
MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

Baby & Child
Research Center

Word Learning from distributional cues

THE [CAT] IS ON THE TABLE
WHERE DOES THE [CAT] SLEEP?
THE [CAT] ATE YOUR FOOD!

- Children can track distributional information at a very young age (Saffran et al. 1996, Aslin et al. 1998, …).

- How does distributional information shape meaning representations and influence word learning?
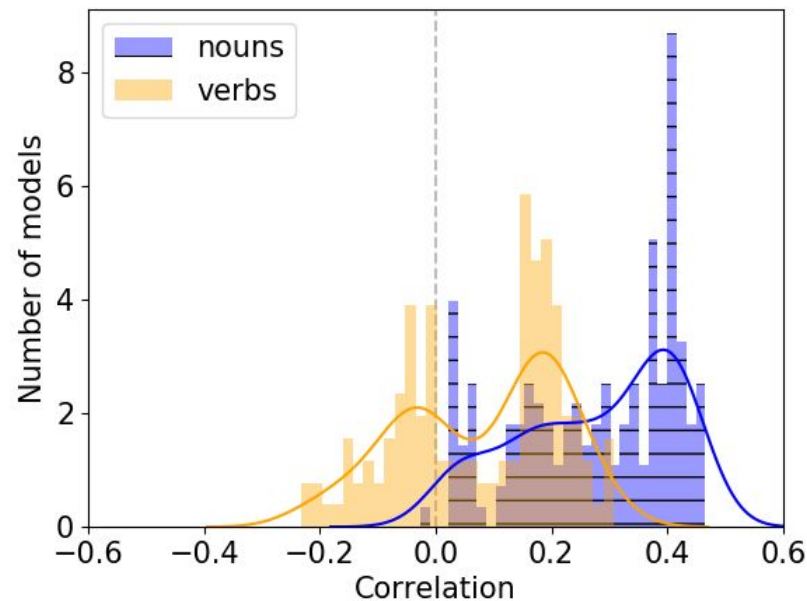
  → We look at Distributional Semantic Models

Method

- Distributional Semantic Models trained on child-directed speech

  - Models: SVD (context-counting) and *word2vec-Skipgram* (context-predicting)

  - Data: CHILDES, age 0 to 60 months

- Evaluation:

  - Correlation between Neighbourhood Density (ND) and Age of Acquisition norms (Wordbank)

M A X
P L A
N C K

Results
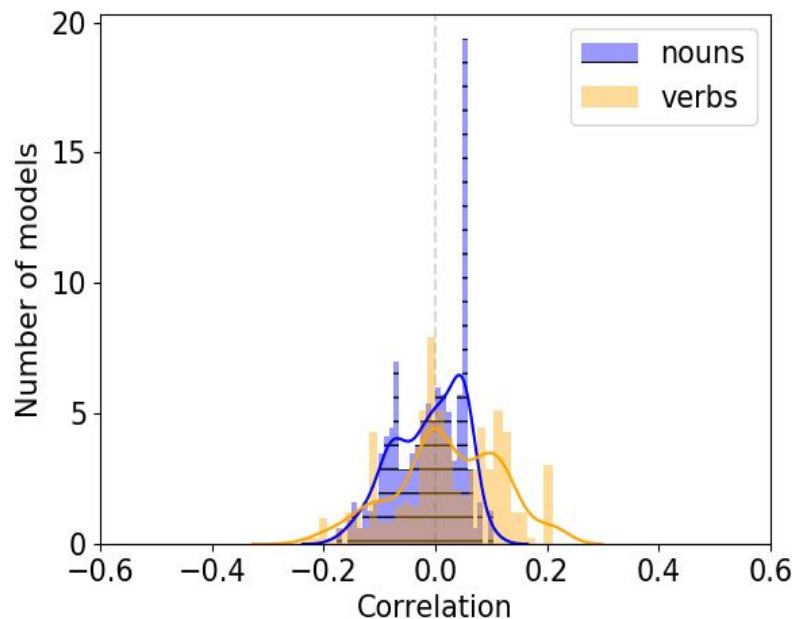
Context-counting   Context-predicting

Findings

1. Context-predicting approach yields a better fit to the data.

2. Negative correlation: words in denser semantic neighbourhoods are associated with delayed acquisition.

3. Parameter analyses: a window size of 1 provides the best fit to the data.

4. Frequency analyses: frequent words tend to have fewer neighbours and are acquired earlier.

Conclusions

◈ Words in denser semantic neighbourhoods are associated with delayed acquisition.

◈ A word's local context is most informative

  ◇ aligns with proposal that children's constrained memories are advantageous in acquisition (Johnson&Newport, 1989).

◈ Frequent words tend to have fewer neighbours

  ◇ tentative explantion for why frequency aids acquisition.