

# A Contextual Analysis of Concrete and Abstract Words



Daniela Naumann, Diego Frassinelli, Sabine Schulte im Walde  
{daniela.naumann, frassinelli, schulte}@ims.uni-stuttgart.de

Institute for Natural Language Processing, University of Stuttgart, Germany

## Motivation

### Problem Setting

- the characteristics of concrete and abstract words play a key role in the debate about meaning representation in the human mind
- concrete words are grounded in the sensory-motor system, while the meaning of abstract words is derived from the activation of concrete words related to them [1]
- Distributional Hypothesis [2]: similar linguistic contexts tend to imply similar word meanings

Thus, we used

- a distributional semantic approach to perform a quantitative investigation of the contexts of concrete and abstract words in order to
- determine distinct contextual cues

## Our Main Hypotheses

- 1) **concrete** and **abstract** words mainly co-occur with **concrete** words
- 2) **concrete** words occur in a limited set of distinct contexts while **abstract** words appear in a broader range of different contexts
- 3) **concrete** words are easier to predict than **abstract** words, due to the greater contextual variability of **abstract** words

## Example Sentences

- a The **football player** kicked the **football** that laid on the **ground** .
- b She was **scared** of **telling** the **truth** .
- c The **manager** was **grateful** for the **tactful behaviour** of his **employee** towards his **assistants**.

## Results

### Hypothesis 1: Co-occurrence

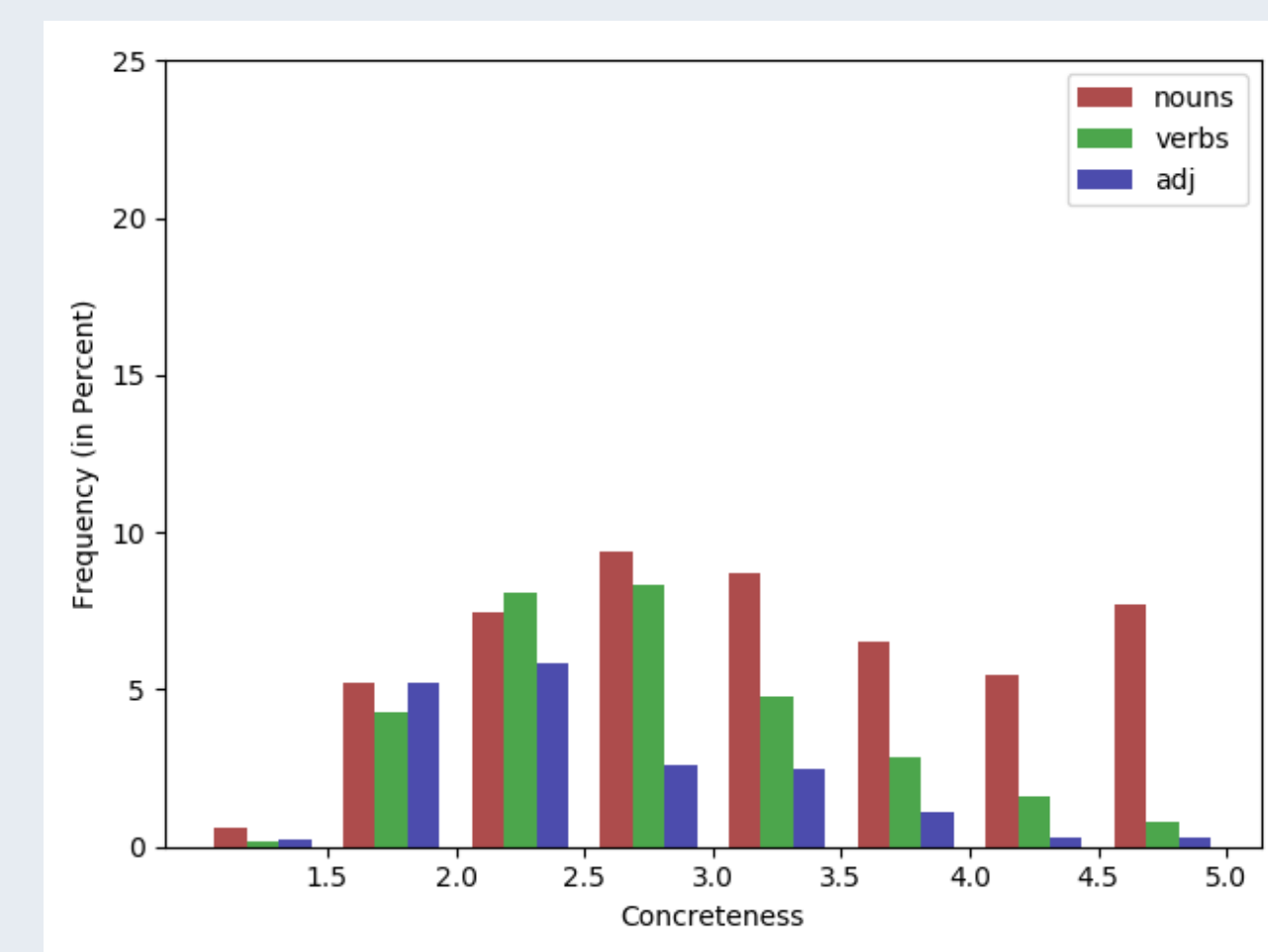


Figure 1: Context of Abstract Nouns

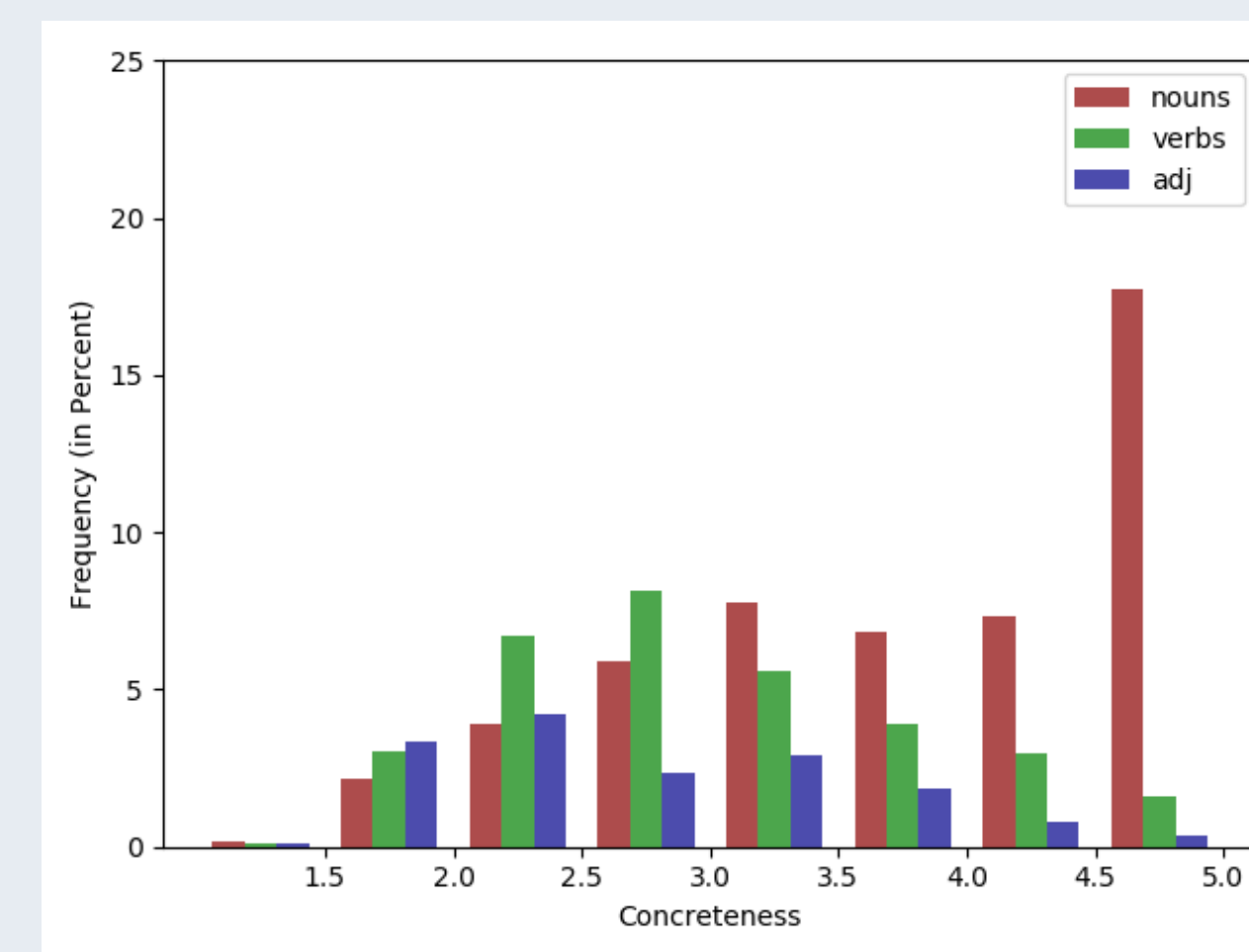


Figure 2: Context of Concrete Nouns

context **concrete** nouns:

- portion conc. nouns > portion abst. nouns

context **abstract** nouns:

- portion conc. nouns < portion abst. nouns

### Hypothesis 2: Semantic Variation of Context

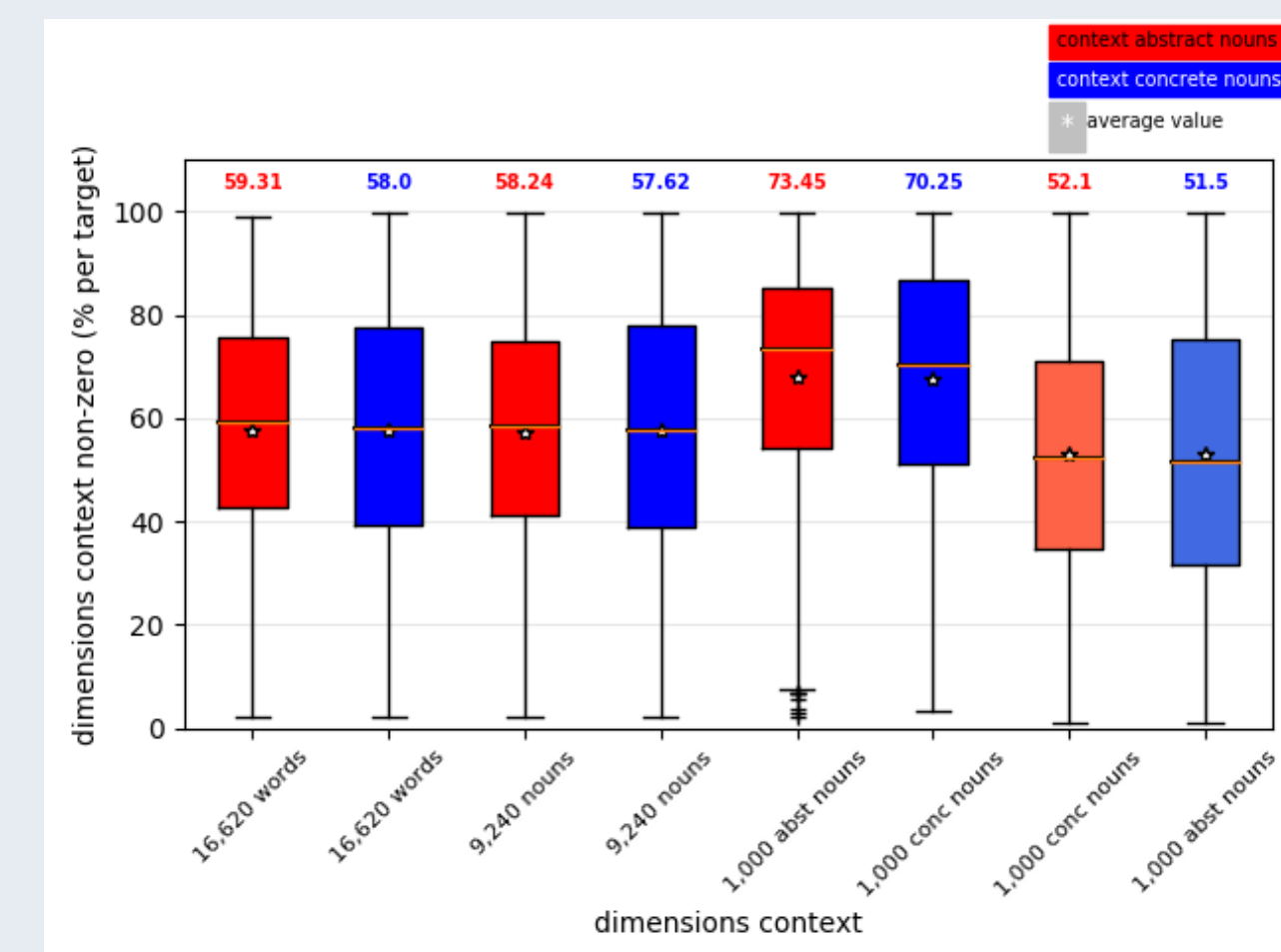


Figure 3: Non-Zero Context Dimensions

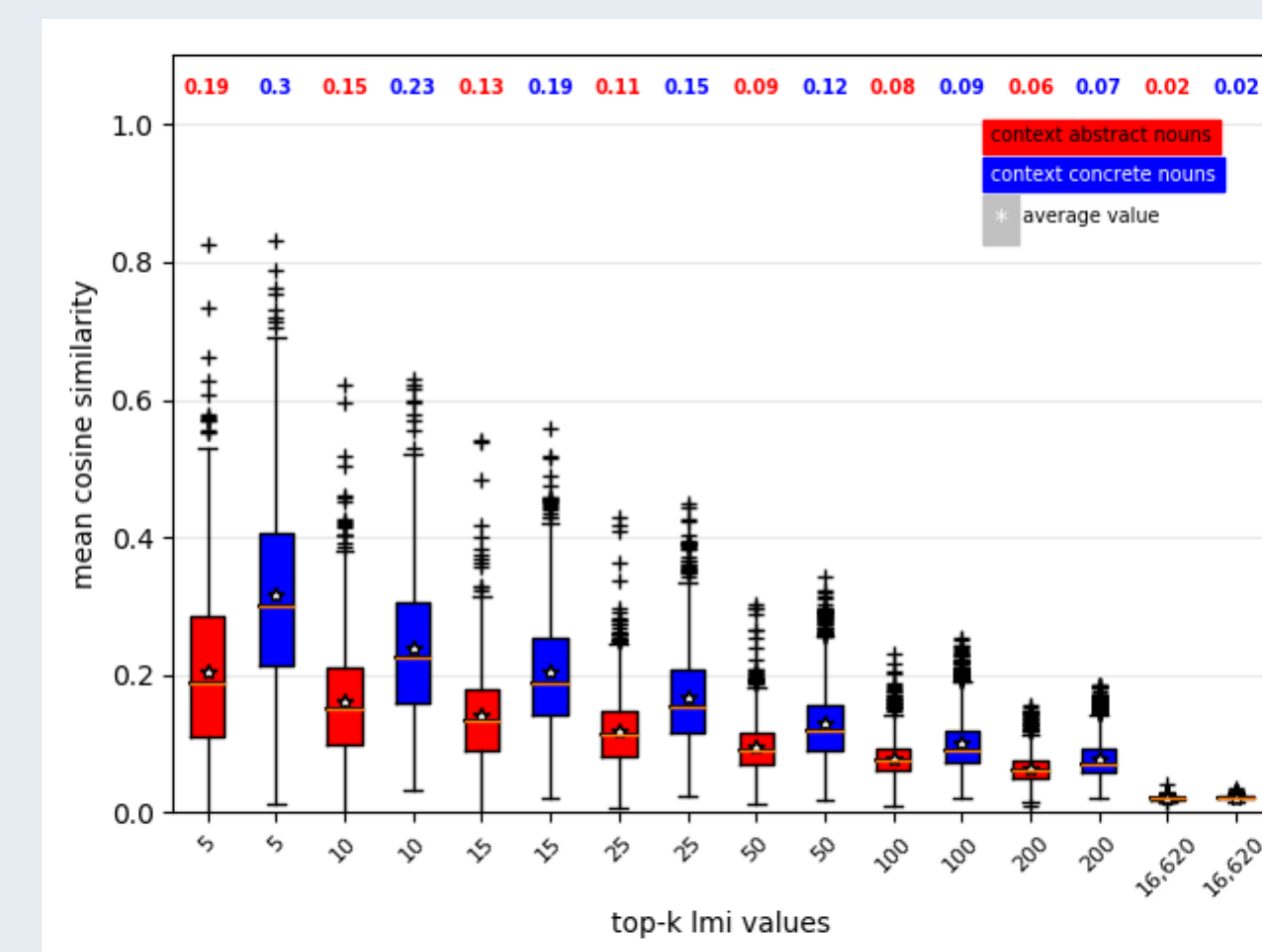


Figure 4: Mean Pairwise Cosine Sim.

context **concrete** nouns:

- high number of contexts
- high context similarity

context **abstract** nouns:

- high number of contexts
- low context similarity

### Hypothesis 3: Contextual Entropy

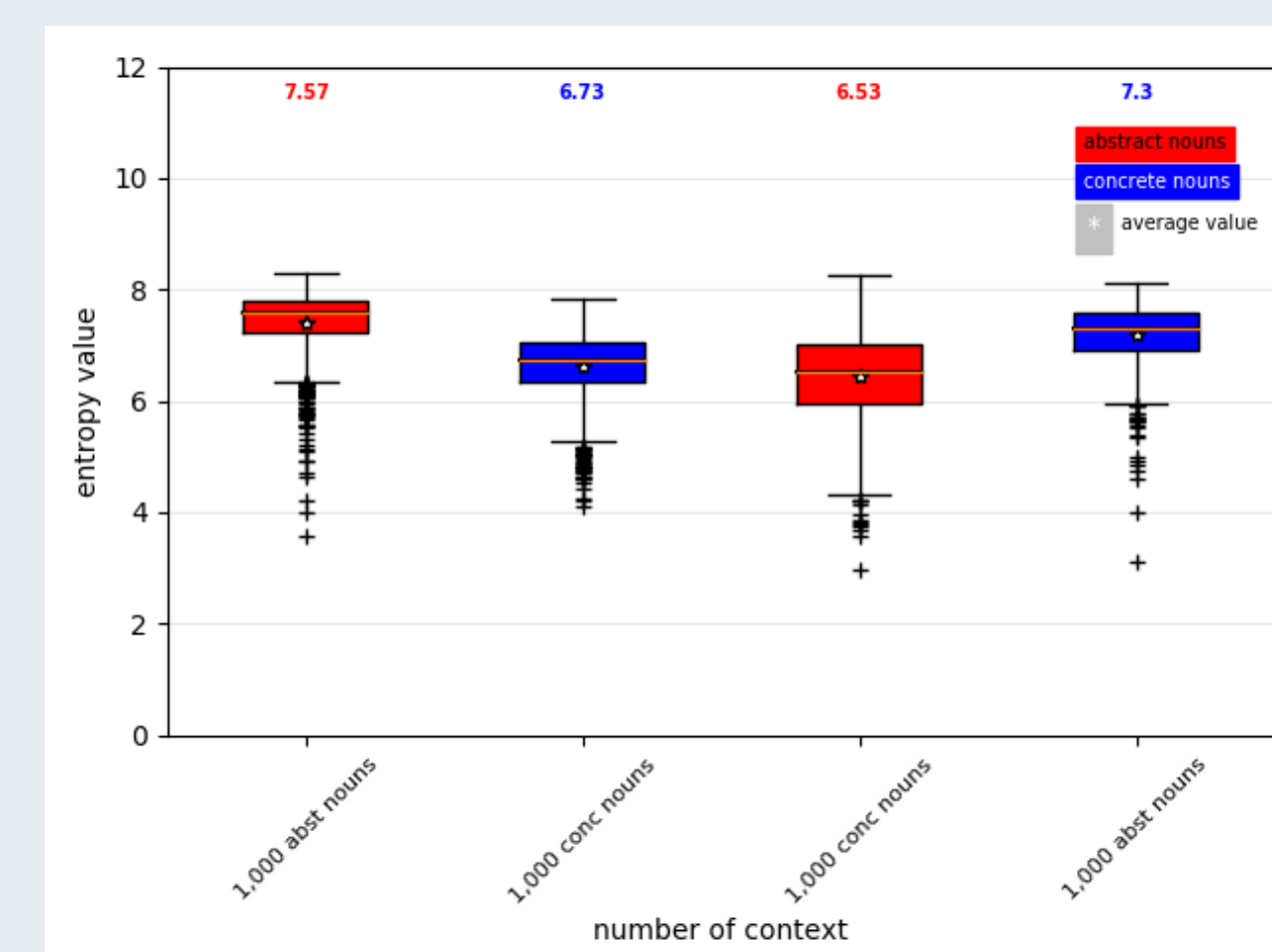


Figure 5: Entropy of Nouns

context **concrete** nouns:

- low contextual entropy value

context **abstract** nouns:

- high contextual entropy value

## Data

### Corpus

- ENCOW16AX (sentence-shuffled) [3]

### Ratings

- Concreteness Ratings [4]

## Conclusion

Our studies show consistent differences in the contexts of concrete and abstract words:

- **concrete** words primarily co-occur with other **concrete** words
- the contexts of **concrete** words can be predicted with greater certainty than the contexts of **abstract** words
- **abstract** words mainly co-occur with other **abstract** words

For abstract words, we yield patterns that challenge the grounding theory of cognition and therefore require further investigation.

## References

- [1] Lawrence W Barsalou and Katja Wiemer-Hastings. Situating abstract concepts. *Grounding cognition: The role of perception and action in memory, language, and thought*, pages 129–163, 2005.
- [2] John R Firth. A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*, 1957.
- [3] Roland Schäfer and Felix Bildhauer. Building large corpora from the web using a new efficient tool chain. In *LREC*, pages 486–493, 2012.
- [4] Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior research methods*, 46(3):904–911, 2014.