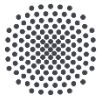
Institute for Natural Language Processing



University of Stuttgart

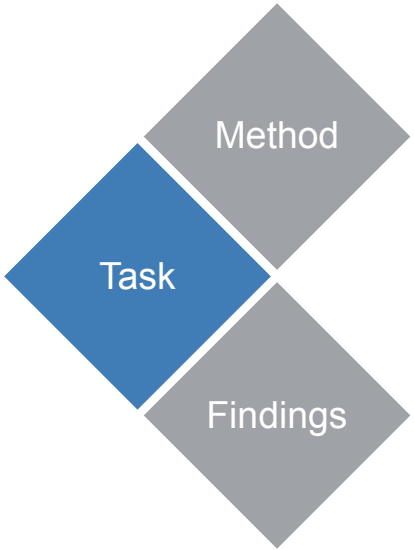
More than just Frequency? Demasking Unsupervised Hypernymy Prediction Methods

Thomas Bott, Dominik Schlechtweg & Sabine Schulte im Walde

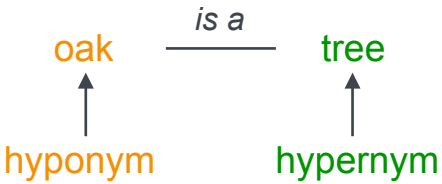
Findings of ACL 2021

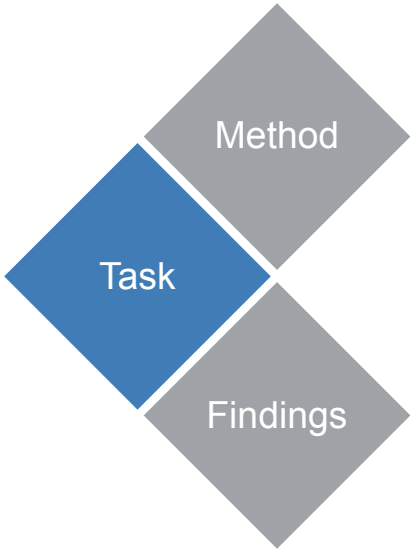
- hypernymy prediction

oak —*is a*— tree

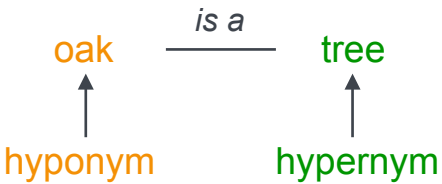


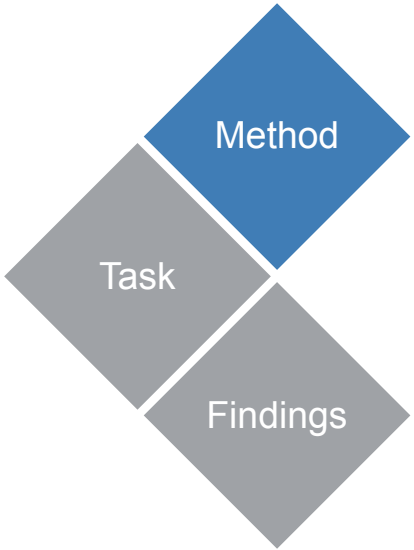
- hypernymy prediction





- hypernymy prediction



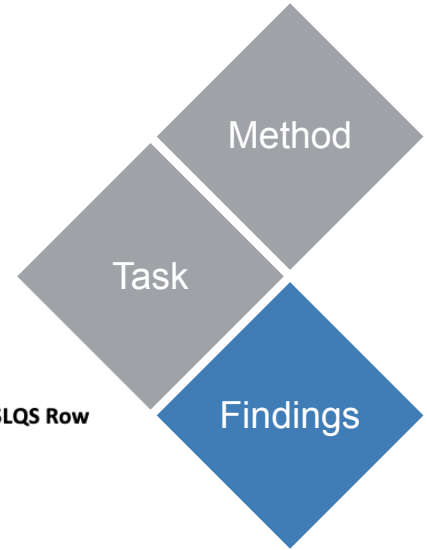


- 4 methods
  - *WeedsPrec, invCL, SLQS Row, SLQS Sec*
- 2 baselines
  - *Word Length, Word Frequency*
- traditional count vector spaces
- WaCky corpora: English and German
- 2 WordNet datasets: English and German
- 4 benchmark datasets (English)
  - *BLESS, EVALution, Lenci/Benotto, Weeds*

- hypernymy prediction

	Length	Freq	WeedsPrec	InvCL	SLQS Row
Freq	0.592				
WeedsPrec	0.604	0.947			
InvCL	0.592	0.992	0.951		
SLQS Row	0.599	0.914	0.938	0.912	
SLQS Sec	0.547	0.715	0.719	0.714	0.727

Correlations between methods



- 3 methods (*WeedsPrec, invCL, SLQS Row*) highly correlated with frequency-based predictions
- *SLQS Sec* shows lower accuracy, but makes correct predictions where the others go wrong

➡ general need to check the frequency bias of a computational method in order to identify frequency-(un)related effects

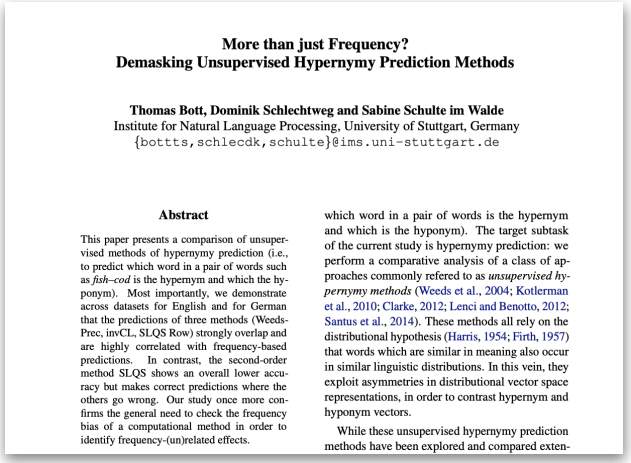
Thank you!

Thomas Bott, Dominik Schlechtweg & Sabine Schulte im Walde

Institute for Natural Language Processing

University of Stuttgart, Germany

{bottts, schlecdk, schulte}@ims.uni-stuttgart.de



More than just Frequency?
Demasking Unsupervised Hypernymy Prediction Methods

Thomas Bott, Dominik Schlechtweg and Sabine Schulte im Walde
Institute for Natural Language Processing, University of Stuttgart, Germany
{bottts,schlecdk,schulte}@ims.uni-stuttgart.de

Abstract

This paper presents a comparison of unsupervised methods of hypernymy prediction (i.e., to predict which word in a pair of words such as *fish–cod* is the hypernym and which the hyponym). Most importantly, we demonstrate across datasets for English and for German that the predictions of three methods (WeedsPrec, invCL, SLQS Row) strongly overlap and are highly correlated with frequency-based predictions. In contrast, the second-order method SLQS shows an overall lower accuracy but makes correct predictions where the others go wrong. Our study once more confirms the general need to check the frequency bias of a computational method in order to identify frequency-(un)related effects.

which word in a pair of words is the hypernym and which is the hyponym). The target subtask of the current study is hypernymy prediction: we perform a comparative analysis of a class of approaches commonly refered to as *unsupervised hypernymy methods* (Weeds et al., 2004; Kotlerman et al., 2010; Clarke, 2012; Lenci and Benotto, 2012; Santus et al., 2014). These methods all rely on the distributional hypothesis (Harris, 1954; Firth, 1957) that words which are similar in meaning also occur in similar linguistic distributions. In this vein, they exploit asymmetries in distributional vector space representations, in order to contrast hypernym and hyponym vectors.

While these unsupervised hypernymy prediction methods have been explored and compared exten-

Findings of ACL 2021