

Resource-poor Translation of Multiword Expressions



SHIVA TASLIMIPOOR & RUSLAN MITKOV

UNIVERSITY OF WOLVERHAMPTON

Relieving the Pain in the Neck in Natural Language Processing:
PARSEME 7th final general meeting, 26-27 September 2016,
Dubrovnik, Croatia

Objectives



- Help linguists to cross-lingually investigate multiword collocations
 - Expressions to be listed in bilingual dictionaries
- Improve extracting multiword collocations in one language using the strength of the other language in collocation identification
- Find translation equivalents for collocations using comparable corpora
 - Using the state-of-the-art distributional similarity approach: word2vec
 - Extracting bilingual contexts from comparable corpora

Findings



- Bilingual context construction is effective for the task of translation equivalent learning for collocations
- Our method outperforms a simplified distributional similarity baseline in finding translation equivalents of verb-noun collocations.
- Our method which uses word2vec to learn distributional similarity is less vulnerable to noise in the corpus.