

# Detecting Multiword Expressions by Dependency Parsing

István Nagy T. and Veronika Vincze

University of Szeged, Hungary  
Department of Informatics



# Introduction

- Automatic detection of MWEs by dependency parsers in different languages
- 3 settings:
  - English verb-particle constructions
  - Hungarian light verb constructions
  - German light verb constructions



# Methods

## English VPCs

Penn Treebank has VPC annotation

Bohnet and Stanford parsers were trained on it

Evaluated the parsers on Wiki50 corpus manually annotated for VPCs

## German LVCs

TIGER corpus has LVC annotation

Bohnet parser trained on TIGER

Evaluated this model on the German part of JRC-Acquis manually annotated for LVCs

# Methods & Results

## Hungarian LVCs

LVCs were manually annotated in the Szeged Treebank

LVC specific dependency relations

Trained and evaluated the Bohnet parser in 10-fold manner

## Results:

The Bohnet parser performs well on all the three tasks

