

## PARSEME

**Parsing and Multi-Word Expressions.** Towards linguistic precision and computational efficiency in Natural Language Processing

### Objectives

- To put **multilingualism** in focus of linguistic and technological studies.
- To establish a long-lasting collaboration of **Natural Language Processing (NLP)** experts within a **cross-lingual, cross-theoretical** and **cross-methodological** research network.
- To bridge the gap between **linguistic precision** and **computational efficiency** in NLP applications.

### Working Groups

- WG 1: **Lexicon/Grammar Interface**
- WG 2: **Parsing Techniques for Multi-Word Expressions**
- WG 3: **Statistical, Hybrid and Multilingual Processing of Multi-Word Expressions**
- WG 4: **Annotating Multi-Word Expressions in Treebanks**

### Main Achievements

- Gathering a cross-lingual, cross-theoretical and cross-methodological consortium.
- 30 countries, 29 languages, 6 dialects and 9 language families are represented.
- Structuring the action among 4 working groups, strong links among groups.
- Initiating highly contrastive review of the state-of-the-art methods taking the variety of languages into account
- 9 short-term visits of, especially young, researchers in partner countries
- 4 national spin-off project proposals
- Strong links with a new European FP7 project
- Strong links with an established workshop

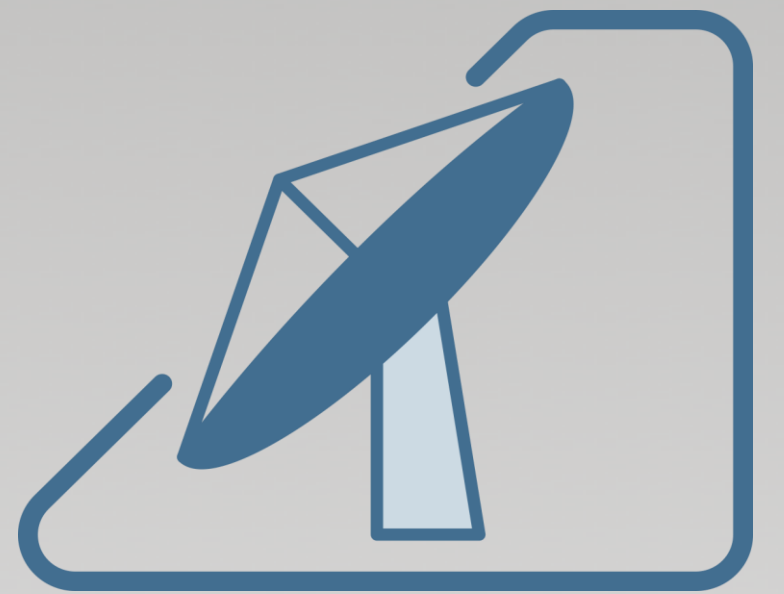
### Gender Balance and Early Stage Researchers

- **50%** of initiators, **40%** of the MC and **54%** of the Steering Committee are **women**
- **50%** of initiators, **21%** of the MC and **46%** of the Steering Committee are **Early Stage Researchers**

### Dissemination

- **Who?**
  - Action's members, young researchers, Master and PhD students, related Actions, projects and networks, other NLP researchers, language industry professionals
- **What?**
  - Action's public and private website, mailing and discussion lists, short term missions, Multi-Word Expressions workshop co-located with the EACL 2014 conference, scientific publications.
- **How?**
  - Presentation slides, abstracts, common papers and short mission reports on the website
  - Workshop proceedings
  - Early-stage researchers as website developers and web masters
  - Regular updates of the mailing list membership

Information and Communication Technologies (ICT)



Participating countries: 30

BG, CH, CZ, DE, DK, EE, EL, ES, FR, HR, HU, IE, IL, IS, IT, LV, LT, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, TR, UK

Internat. Collaboration:

BR, US

### Contact details

#### Chair of the Action

Agata Savary  
agata.savary@univ-tours.fr

#### Domain Committee Rapporteur

Roel J Van den Berg  
rvandenberg@ese.eur.nl

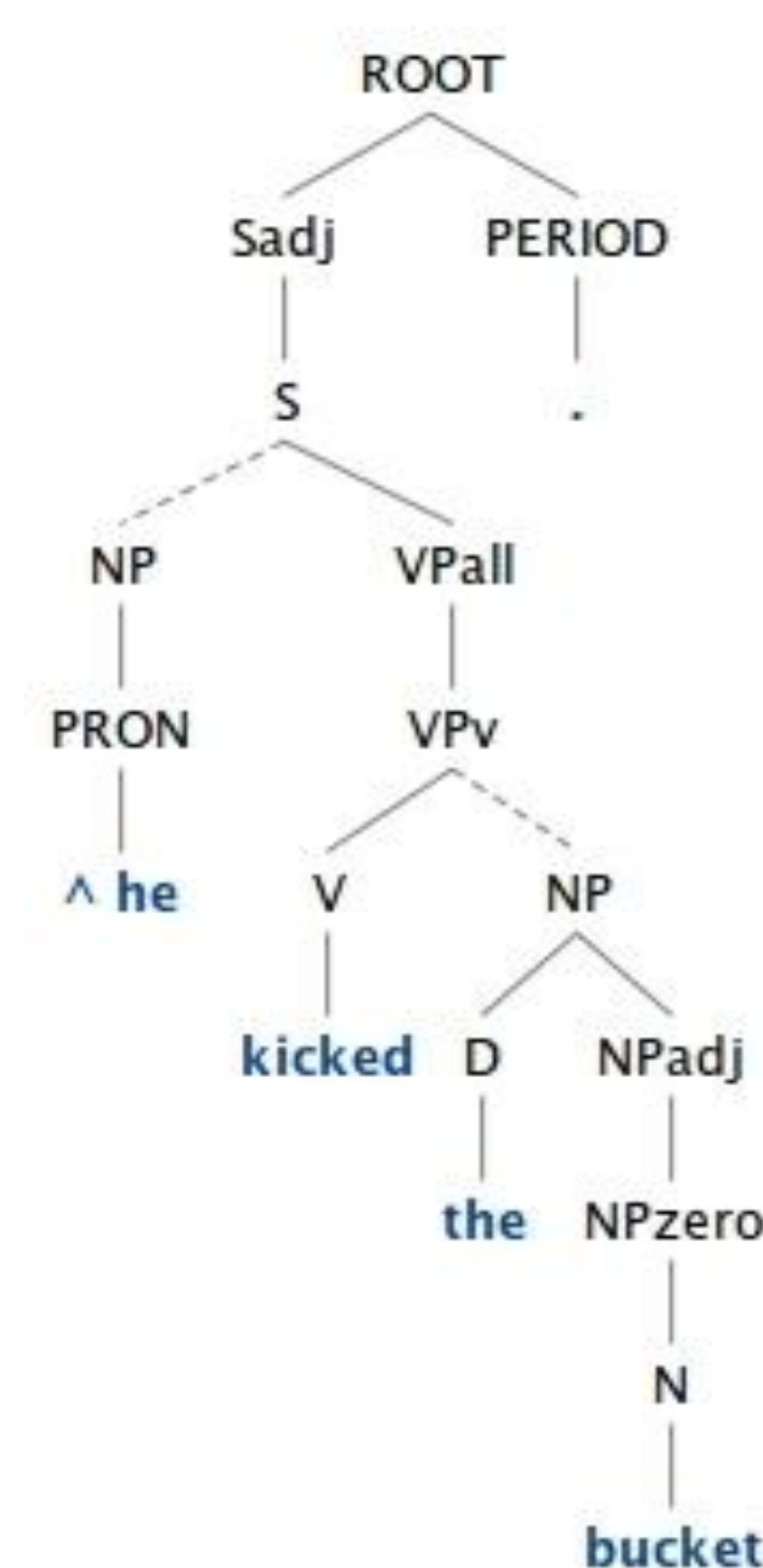
#### Science Officer (COST Office)

Ralph Stübner  
Ralph.Stuebner@cost.eu

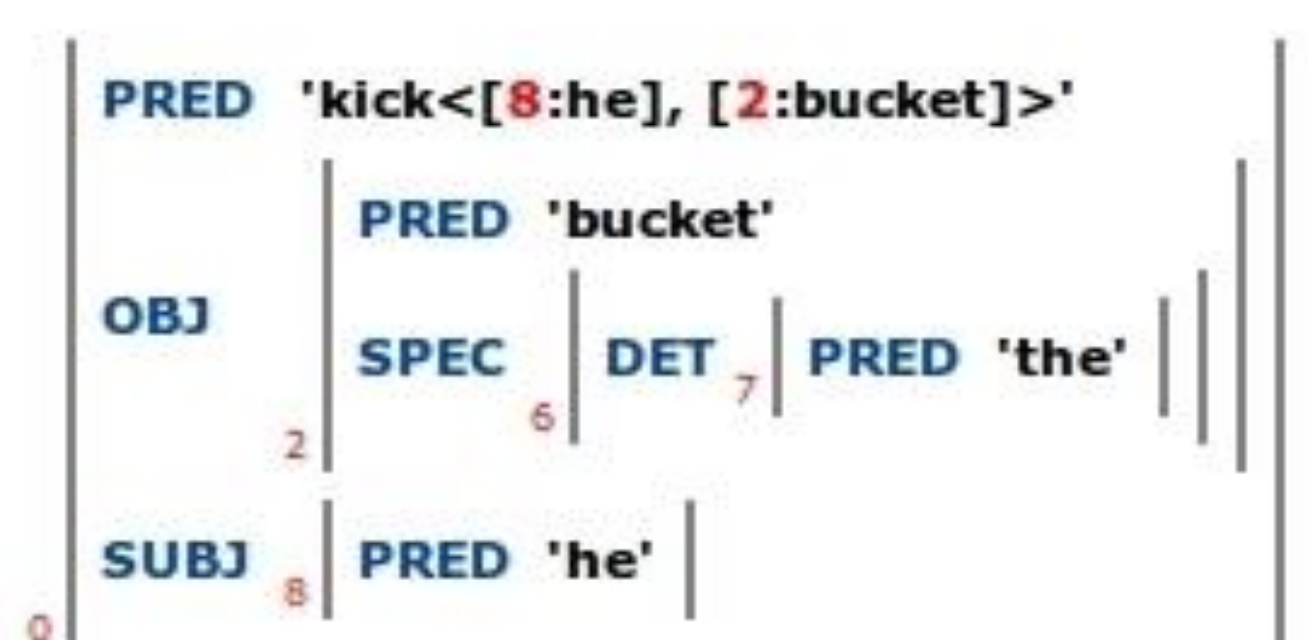
#### Website

[www.parseme.eu](http://www.parseme.eu)

#### C-structure



#### F-structure



Sample parsing structures for the idiom "to kick the bucket"



COST is supported by the EU RTD Framework Programme



ESF provides the COST Office through a European Commission contract