

COST Action no. IC1207

2013 | 2017

# 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: **Hybrid Parsing of Multi-Word Expressions**
- WG 4: **Annotating Multi-Word Expressions in Treebanks**

## Main Achievements

- Gathering a cross-lingual, cross-theoretical and cross-methodological **consortium**.
- 23 languages, 6 dialects and 8 language families are represented: **Germanic** (British/American English, Danish, Dutch, German, Icelandic, Norwegian, Swedish), **Finno-Ugric** (Estonian, Hungarian), **Hellenic** (Greek), **Romance** (Swiss /France French, Italian, European/Brazilian Portuguese, Spanish), **Semitic** (Hebrew, Maltese), **Slavic** (Bulgarian, Czech, Polish, Serbian, Macedonian), **Turkic** (Turkish).

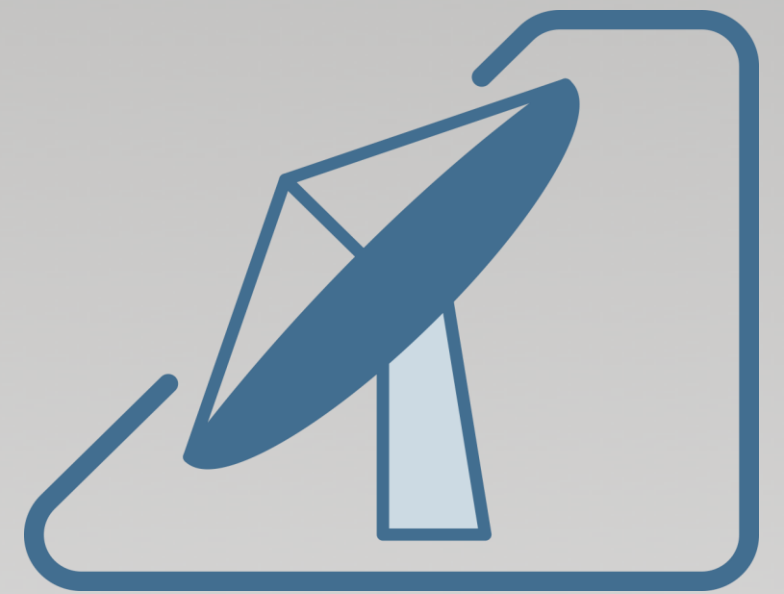
## Gender Balance and Early Stage Researchers

- Objectives:
  - Respect an appropriate gender balance in all activities.
  - Promote a good integration of ESRs in the ERA.
- Status:
  - **50%** of initiators, **40%** of the MC and **44%** of the SC are **women**
  - **50%** of initiators, **30%** of the MC and **22%** of the SC are **ESRs**
  - The role of the **Representative of ESRs** has been defined within the Steering Committee
- Foreseen Support Measures: STSMs for ESRs, Training Schools, Workshops

## Dissemination Plan

- Who?
  - Action's members, ESRs, Master and PhD students, related Actions (MUMIA, IS1006), projects and networks, other NLP researchers, language industry professionals, language resource providers, European institutions, foreign language teachers and learners, general public
- What?
  - Action's public and private website, mailing and discussion lists, training schools, STSMs, workshops in established conferences, scientific publications, Master's and doctoral training, European infrastructures for language resources (META-SHARE), press releases
- How?
  - Dissemination Manager nomination
  - Annual evaluation by the SC of the dissemination plan fulfillment.

Information and Communication Technologies (ICT)



Participating countries: 24

CH, CZ, DE, DK, EE, EL, ES, FR, HR, HU, IL, IS, IT, MK, MT, NL, NO, PL, PT, RS, SI (intention), SE, 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

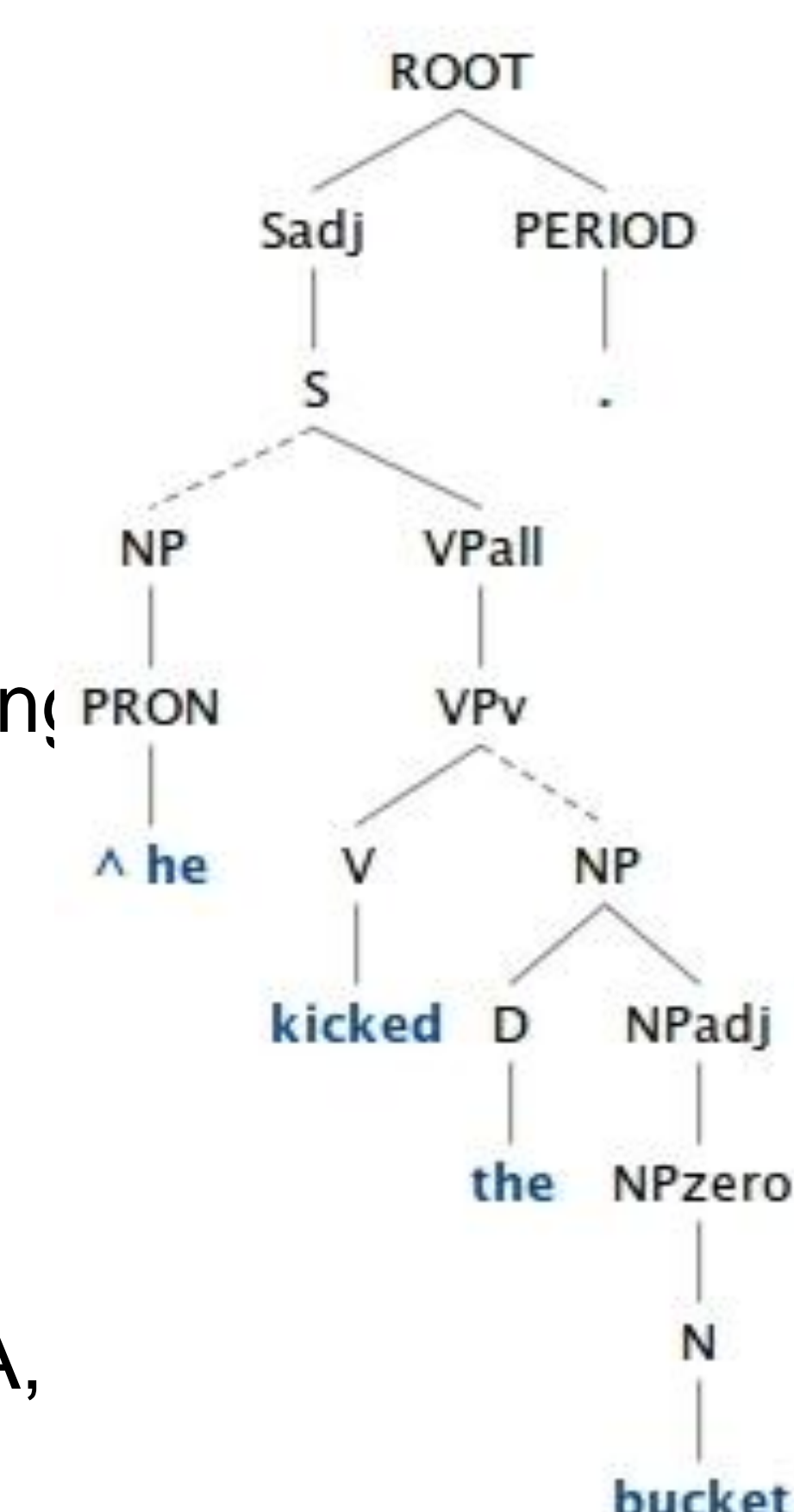
Science Officer (COST Office)

Ralph Stübner  
Ralph.Stuebner@cost.eu

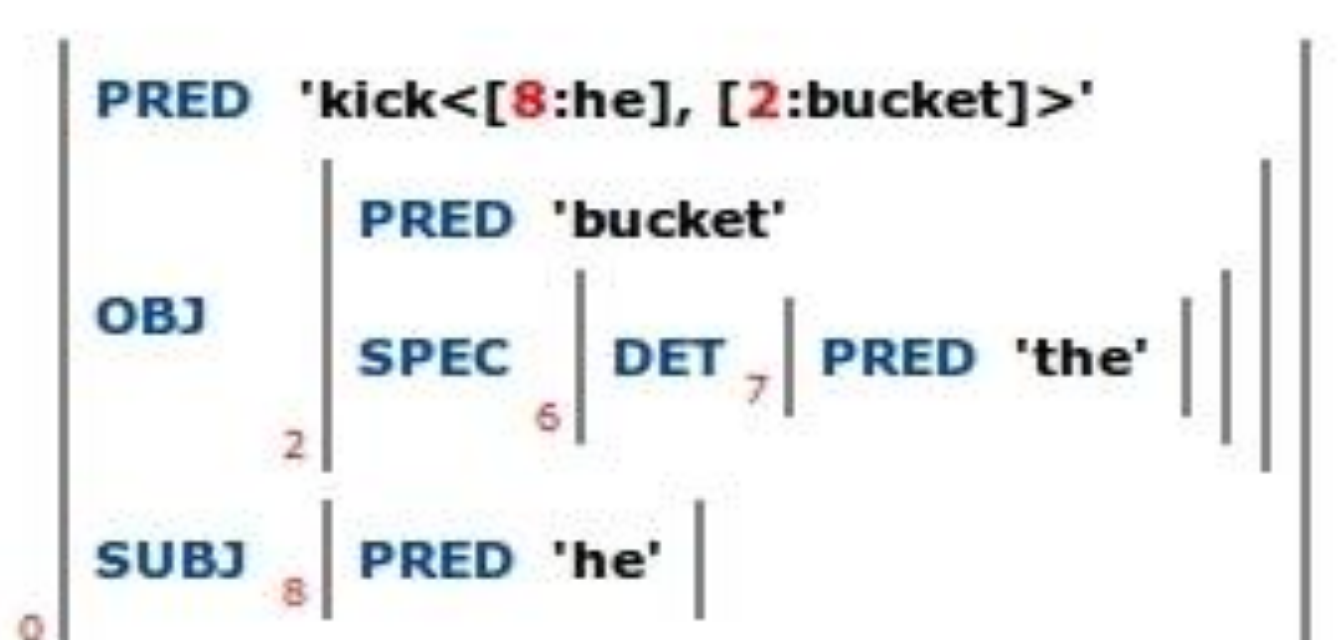
Website

[www.cost.eu/domains\\_actions/ict/Actions/IC1207](http://www.cost.eu/domains_actions/ict/Actions/IC1207)

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