

Working Group 3: Multilingual and Cross-Lingual Language Technology

- **Leader:** [Joakim Nivre](#) (Sweden)
- **Vice-leader:** [Gülşen Cebiroğlu Eryiğit](#) (Turkey)

Workplan

Unified modelling helps solve NLP tasks with higher accuracy and better awareness of diversity. Therefore, this WG will be dedicated to NLP coordinating the development of tools leveraging universality and promoting diversity:

1. Multilingual and cross-lingual **syntactic parsers** which:
 - pay attention to hard and underrepresented phenomena (unbounded dependencies, MWEs,...),
 - leverage transfer of annotations or models in order to cope with data scarceness;
2. Prototypes of multilingual and cross-lingual **semantic parsers** which:
 - derive bi-lexical semantic dependencies from syntactic trees,
 - resolve idiosyncrasies in the syntax-semantics interface;
3. Multilingual **MWE discovery** tools which:
 - exploit large non-annotated data to compensate the sparseness of MWEs in annotated corpora,
 - are coupled both with lexicons and MWE identifiers;
4. Multilingual **MWE identifiers** which:
 - are coupled with MWE discovery and lexica to better handle unseen data,
 - pay attention to underrepresented phenomena, e.g., discontinuity/variability of MWEs;
5. Prototypes of tools for automatic identification of idiosyncratic **constructions**.

The tools themselves will be funded at the national level. WG3 will bring the federating effect to these activities, notably by organizing multilingual **evaluation campaigns** on parsing and MWE identification. Diversity-based evaluation measures from WG4 will be promoted. The outcomes should validate the computational tractability of the terminologies unified in WG1.

Members and organisation

- [List](#) of current WG3 members
- Activities are currently structured around four primary [WG3 tasks](#) detailed below, but proposals for new activities are always welcome.

Upcoming meetings

- WG Meeting 9 (online) - 10 June 2024, 10:00 **CEST**

Minutes of past meetings

- [WG3 Meeting 1 Minutes](#) **16-17 March 2023**, Paris-Saclay University, France (co-located with UniDive 1st general meeting)
- [WG3 Meeting 2 Minutes](#) **8 September 2023**, Istanbul Technical University, Türkiye
- [WG3 Meeting 3 Minutes](#) **20 November 2023**, online
- [WG3 Meeting 4 Minutes](#) **18 December 2023**, online
- [WG3 Meeting 5 Minutes](#) **15 January 2024**, online
- [WG3 Meeting 6 \(including joint meetings with WG1 and WG4\)](#) **9 February 2024**, University of Naples L'Orientale, Italy
- [WG3 Meeting 7 Minutes](#) **11 March 2024**, online
- [WG3 Meeting 8 Minutes](#) **29 April 2024**, online

WG3 Tasks

- **Task 3.1 Documentation of multilingual tools and resources**
 - Leaders / Contacts: A. Seza Doğruöz, Teresa Lynn, Maria Giagkou
 - Objectives: Assessing the “discoverability” of NLP tools and resources, and analyzing the NLP tool availability in the ELG catalogue.
 - Workplan: Data Collection and Analysis.
 - How can I contribute: Please fill out the shared [document](#) for your languages in focus by following the task description in slide three of the following presentation
.
 - Documents / Links
task description
- **Task 3.2 Evaluation campaign: morphosyntactic parsing**
 - Leaders / Contacts: Omer Goldman, Leonie Weissweiler, Reut Tsarfaty
 - Objectives: Organization of the first WG3 evaluation campaign on morphosyntactic parsing, which aims to combine syntactic parsing with morphological analysis in a way that avoids (most) theoretical debates on word boundaries.
 - Workplan: Data preparation in 2024, shared task in 2025
 - How can I contribute: You may contribute for data preparation and join the campaign by writing to: msap-discussion-group@googlegroups.com
 - Documents / Links ([abstract](#), [slides](#))
- **Task 3.3 Conceptions of multilinguality**
 - Leaders / Contacts: Adriana Pagano (apagano@letras.ufmg.br), Ilan Kernerman (ilan@lexicala.com)
 - Objectives: Define the concepts of *multilingual*, *cross-lingual*, and *translingual* in the context of Language Technology.
 - Workplan: Devise a survey for the Action members and analyze (and publish) the results.
 - How can I contribute: If you like to contribute to the data analysis, please contact the task leaders.
 - Documents / Links: [
survey results
]

Channels

- [WG3 mailing list](#) for general announcements and proposals
- [WG3 Telegram group](#) for special announcements and discussions

From:

<https://unidive.lisn.upsaclay.fr/> - **Universality, diversity and idiosyncrasy
in language technology
CA21167 COST Action**

Permanent link:

<https://unidive.lisn.upsaclay.fr/doku.php?id=wg3:wg3>

Last update: **2024/05/06 12:18**

