

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.

Current Subtasks

- Documentation of multilingual tools and resources [A. Seza Doğruöz, Teresa Lynn, Maria Giagkou]
- Evaluation campaign: morphosyntactic parsing [Omer Goldman, Leonie Weissweiler, Reut Tsarfaty]
- Defining multilingual and cross-lingual tools supporting diversity [Adriana Pagano, Ilan Kernerman]

Documents

- [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

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&rev=1705350604>



Last update: **2024/01/15 21:30**