

# UniDive 2nd general meeting

- **Event title:** UniDive 2nd general meeting
- **Organisation:** University of Naples L'Orientale, Department of Literary, Linguistic and Comparative Studies, UNIOR NLP Research Group, Italy
- **Location:** [Palazzo Mediterraneo](#), University of Naples "L'Orientale" (see also the [Google maps link](#))
- **Dates:** 8-9 February 2024 (co-located with a WG1 meeting on 7 Feb 2024)
- **Local Organizers:** [Johanna Monti](#) and [Maria Pia di Buono](#)



## Program

All sessions take place at the University of Naples "L'Orientale", [Palazzo Mediterraneo](#), rooms 1.1 and 1.5. To know more about the posters, see the [call for abstracts](#).

### Wednesday 7 February: WG1 meeting

- 8:30-9:15 Registration
- 9:15-10:30 [Opening session](#) (room 1.1; [zoom link](#); chairs: Bruno Guillaume, Kaja Dobrovoljc)
  - 9:15-9:45 Presentation of WG1 activities and the meeting programme
  - 9:45-10:30 Plenary talk: [Daniel Zeman](#) (Charles University, Prague) - *UD Guidelines and Validation: How It Works and Evolves* ([zoom link](#))
    - Abstract: In this talk, I will discuss several points of the UD guidelines that are particularly important for people who want to take an active role in data preparation, addition of new languages to UD, or data maintenance. Besides an overview of the crucial points in the UD relations taxonomy, the talk will provide

practical advice about the validation and release procedure. I will also briefly present the evolution of the UD guidelines and infrastructure and provide my view of what can be expected in the future.

▪ Slides:

PDF

- 10:30-11:00 *coffee break*
- 11:00-12:30 Parallel working sessions I
  - T1.3 Extensions and updates to morphosyntactic annotation guidelines ([agenda](#), room 1.1, [zoom link](#), chairs: Atul Kr. Ojha, Hiwa Asadpour)
  - T1.2 Pilot PARSEME annotation of non-verbal MWEs ([agenda](#), room 1.5, [zoom link](#), chairs: Agata Savary, Voula Giouli, Stella Markanotatou, Sara Stymne, Carlos Ramisch)
- 12:30-14:00 *lunch*
- 14:00-15:30 Parallel working sessions II
  - T1.4 Manual annotation tools and file formats ([Slides](#), [agenda](#), room 1.1, ; [zoom link](#), chairs: Bruno Guillaume, Frantisek Forgac, A. Seza Doğruöz, Atul Kr. Ojha)
  - T1.2 Discussion of pilot PARSEME annotation ([agenda](#), room 1.5, [zoom link](#), chairs: Agata Savary, Voula Giouli, Stella Markanotatou, Sara Stymne, Carlos Ramisch)
- 15:30-16:00 *coffee break*
- 16:00-17:30 Parallel working sessions III
  - T1.1 Linguistic typology and multilingual corpus annotation ([agenda](#), room 1.1, [zoom link](#), chairs: André Coneglian, Atul Kr. Ojha, A. Seza Doğruöz, Hiwa Asadpour, Manjola Lumani Zaçellari)
  - T1.2/T1.4 Presentation of the PARSEME pipeline to Language Leaders ([agenda](#), room 1.5, [zoom link](#), chairs: Van Tuan Bui, Agata Savary)

For a quick overview of the sessions, see [WG1 Day program overview](#).

## Thursday 8 February: 2nd UniDive Workshop

- 8:15-8:55 registration and poster hanging for session A
- 8:55-9:00 opening and welcome
- 9:00-10:30 (session 1, [zoom link](#); chair: Alina Wróblewska) invited talks:
  - 9:00-9:45 [François Yvon](#) (Sorbonne Université, CNRS, Inserm) - [Training and evaluating massively multilingual language models](#)
    - Abstract: In this presentation, I will discuss the training and evaluation of massively multilingual language models, capable of handling dozens or even hundreds of languages. After motivating the development of such models, I will draw some lessons learned in the course of developing Glot500, a language model covering 500 languages, and some associated resources. I will notably focus on the challenges raised by “low-resourced” languages, i.e. languages for which the available learning data is often incomplete, highly specialised and also possibly very noisy.
  - 9:45-10:30 [Beata Wójtowicz](#) (University of Warsaw) and [Martin Benjamin](#) (Kamusi Project International) -  
African Languages in the Technological Realm
    - Abstract: We examine the relationship between Africa’s 2000 languages and contemporary language technologies. Our lens focuses on Swahili as a case study, the most widely spoken African language, the one with the most research and technological development, and a representative linguistic model for about 500

other Bantu languages. What has been done, what can be done, and what hasn't been done for Swahili, and what does this say about 100 other languages with rapidly growing populations that are each spoken by millions of people around Africa?

- 10:30-11:00 *coffee break*
- 11:00-11:30 (session 2; onsite only) plenary [poster presentations A](#) (1 minute per poster)  
all slides
- 11:30-13:00 (session 3; onsite only) [poster session A](#)
- 13:00-14:30 *lunch*, poster hanging for session B
- 14:30-15:00 (session 4; onsite only) plenary [poster presentations B](#) (1 minute per poster)  
all slides
- 15:00-16:30 (session 5; onsite only) [poster session B](#)
- 16:30-16:45 *coffee break*
- 16:45-18:00 (session 6; onsite only; room 1.5) Extended Core Group meeting
- Social events (details [below](#)):
  - 18:00 Departure from Palazzo Mediterraneo for the guided tour
  - 18:30 A guided tour (starting from Piazza San Domenico)
  - 19:30 A welcome cocktail

## Friday 9 February: WG sessions

- 9:00-10:15 (session 7) parallel working sessions
  - WG1+WG3 (room 1.1; on-site only; chairs: Bruno Guillaume, Kaja Dobrovoljc, Joakim Nivre, Gülşen Eryiğit) [Minutes](#)
    - 9:00-9:45 Presentation of new morphosyntactic representation for shared task (T3.2) followed by discussion (chair: Omer Goldman)
      - [Presentation](#)
    - 9:45-10:15 Presentations from both WGs of current plans on surveying and documenting tools and resources (T1.4, T3.1) followed by discussion (chair: A. Seza Doğruöz, [slides WG1](#), [slides WG3](#))
  - WG2+WG4 (room 1.5; chairs: Verginica Mititelu, Voula Giouli, Marie-Catherine de Marneffe, Abigail Walsh)
    - 9:00-9:30 Promote diversity wrt. cross-language unification of lexical features (chairs: Kilian Evang and Dan Zeman)
    - 9:30-10:15 Development of MWE lexica: requirements and challenges (chairs: Stella Markantonatou and Ivelina Stoyanova)
- 10:15-10:45 *coffee break*
- 10:45-12:00 (session 8, on-site only) parallel working sessions
  - WG1+WG2 (room 1.1; chairs: Bruno Guillaume, Kaja Dobrovoljc, Verginica Mititelu, Voula Giouli)
    - 10:45-11:25 Harmonizing the definition of a “syntactic word” across languages (chairs: Kilian Evang and Dan Zeman)  
PDF
    - 11:25-12:00 Shared treatment of MWEs of various parts of speech. Case study: nominal MWEs (chair: Voula Giouli, [slides](#))
  - WG3+WG4 (room 1.5; chairs: Joakim Nivre, Gülşen Eryiğit, Marie-Catherine de Marneffe, Abigail Walsh) [Minutes](#)
    - 10:45-11:15 Presentations from both WGs on NLP resource documentation tasks undertaken in WG3 and WG4 (T3.1, T4.1) (chair: Lucia Amoros Poveda)

- 11:15-12:00 Short presentation of (a) shared task (T3.2) by WG3 and (b) metrics by WG4, followed by discussions (chair: Joakim Nivre)
  - [Shared task presentation](#) (Omer Goldman)
  - [Diversity metrics for the shared task](#) (Louis Estève)
- 12:00-12:15 *group photo* [1](#)

2  
3
- 12:15-13:30 *lunch*
- 13:30-14:45 (session 9) parallel working sessions
  - WG1 (room 1.1; chairs: Bruno Guillaume, Kaja Dobrovoljc)
    - Reporting the outcomes of February 7 sessions and discussing next steps ([slides](#))
  - WG3 (room 1.4 and room 1.5; [zoom link](#); chairs: Joakim Nivre, Gülşen Eryiğit)
    - Hands-on training for the multilingual tools documentation subtask (room 1.5, chair: Teresa Lynn)
      - [Presentation](#)
      - [Task 3.1.1 data collection form](#)
      - [Task 3.1.2 volunteer form](#)
    - Final discussion of specifications for the shared task on morphosyntactic parsing (room 1.4, chair: Omer Goldman)
- 14:45-15:15 *coffee break*
- 15:15-16:30 (session 10) parallel working sessions
  - WG2 (room 1.1; chairs: Verginica Mititelu, Voula Giouli)
    - 15:15-15:40 Task 2.1.2: Harmonizing lemmatization rules (for words and MWEs) (and standardizing lists of lexemes for auxiliaries, pronouns and determiners) (chairs: Dan Zeman and Kilian Evang)
    - 15:40-16:00 Task 2.2: Demonstration of the process of extending the ELEXIS-WSD corpus with additional annotation layers and other languages (chairs: Jaka Cibej and Carole Tiberius)
    - 16:00-16:20 Task 2.3:
 

Standardizing lexica of MWEs

 (chair: Christian Chiarcos)
    - 16:20-16:30 Task 2.3: Survey on existing MWEs lexica: updates (chairs: Stella Markantonatou and Ivelina Stoyanova)
  - WG4 (room 1.5; chair: Marie-Catherine de Marneffe)
    - 15:15 - 15:45 Task 4.2: Discussion on the properties of diversity functions (Louis Estève)
    - 15:45 - 16:15 Task 4.1:
 

Low Resourced Languages

 . State of the Art (Lucia Amoros Poveda)
    - 16:15 - 16:30 Community discussion of next steps
- 16:30-16:45 *break* (WG leaders prepare summaries)
- 16:45-17:30 (session 11; room 1.1) plenary closing session (summary of WG sessions)

## Registration

TBA

## Directions

The event will take place at: Università degli Studi di Napoli L'Orientale [Palazzo Mediterraneo](#) – Via Nuova Marina, 59, 80133 Napoli

How to get here:

- From the train station Napoli Centrale:
  - It takes approximately 20 minutes from the train station Napoli Centrale to Palazzo del Mediterraneo.
  - By underground railway: Metro Linea 1 to Piscinola – Università or Duomo stop.
- How to get by car:
  - From A1/E45 Take the exit toward Napoli Centro/Porto.
  - Merge onto A3.
  - Take the exit toward Centro.
  - Continue on Via Reggia di Portici.
  - Take Via Amerigo Vespucci to Via Nuova Marina.

## Accommodation

Some recommended hotels are marked in [this map](#).

Please note that Naples is a highly crowded tourist destination even during non-holiday periods. There are many bed and breakfasts in the area adjacent to the university, but we recommend booking accommodation as early as possible.

## Meals

TBA

## Social Events

We plan to have two social events on the 8th of February. The first event is a **guided walking tour of the city center**. The tour will last approximately one hour and will showcase some of the most iconic places in the city.

**Guided walking tour update** We will move from **Palazzo del Mediterraneo** at **18:00** (not at 18:30) and reach the tour starting point in **\*\*Piazza San Domenico - Palazzo Corigliano\*\*** to meet our guides. We will move from there at **18:30** for a short (1 hour) guided tour in the city center. Based on the number of participants, we will organize groups of about 30 people.

The second event is a **welcome aperitif** that will take place at [Lazzarelle Bistrot](#) near the National Archaeological Museum. Those who will not participate in the guided tour can reach the venue where the aperitif will be held at 7:30 PM ([25 mins on foot, 17 mins by public transportation](#)).

## Reimbursement

If you are selected by the Core Group to be reimbursement for the meeting, please accept the invitation that will be sent to you via e-COST before the meeting. After the meeting, fill in the reimbursement form that will be available online via the e-COST system and upload all relevant signed documents, as indicated in the reimbursement form. Before making any arrangements PLEASE READ general rules for the reimbursement described in the [COST Annotated Rules](#), Annex A1-3.1 TRAVEL REIMBURSEMENT RULES. External members (who have not received the e-COST invitation) can also participate freely but will not be reimbursed for their travel and stay. Those external members should also fill in the registration form (to be published soon).

The daily reimbursement rate for this meeting is 160 EUR per night (flat rate). The number of the reimbursed nights is equal to the number of attended workshop/meeting days plus 1 (maximum one night before and one night after the meeting).

## Expression of interest

If you are willing to participate in the Naples meeting, please, fill in the [expression of interest](#) form.

Please, note that filling in this form does not ensure you an official invite to the meeting or reimbursement of your travel and stay. Decisions about reimbursements will be made starting from late December.

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=meetings:general\\_meetings:2nd\\_unidive\\_general\\_meeting&rev=1708421568](https://unidive.lisn.upsaclay.fr/doku.php?id=meetings:general_meetings:2nd_unidive_general_meeting&rev=1708421568)

Last update: **2024/02/20 10:32**

