Report on the activities of the SIGLEX-MWE Section in 2024

Voula Giouli, Verginica Barbu Mititelu, Kilian Evang, Lifeng Han, A. Seza Doğruöz, Alexandre Rademaker

January 2025

According to the constitution of the [SIGLEX-MWE Section](), the standing committee should report on the Section's activities to its members yearly. This document is the report on the Section's activities in 2024.

SIGLEX elections — Section representative

Following the [SIGLEX constitution](), the term of the SIGLEX board members elected in 2022, including the MWE section representative, finished in 2024, requiring the organization of elections. A call for nominations was sent to all SIGLEX members on March 11, 2024, and was open until May 1st, 2024. Steven Bethard (SIGLEX president) and Voula Giouli (MWE Section representative) collected biographies and candidacy statements. ACL provided the technical infrastructure for the online elections on the ElectionRunner platform.

Two nominations were received for the position of MWE Section representative, namely Verginica Barbu Mititelu (RACAI, Romania) and Alexandre Rademaker (IBM Research and EMAp/FGV, Brazil). All 501 SIGLEX members registered until July 15, 2024, received an individual ballot email and 112 SIGLEX members voted. The elections lasted one month (July 22 - August 21, 2024). Verginica Barbu Mititelu received 70% of the votes while Alexandre Rademaker received 17%, and 11% abstained. Verginica Barbu Mititelu was elected thus becoming the MWE Section representative for the term 2024-2026.

The other standing committee members are not elected but nominated (see below).

Standing committee

Since August 2018, the SIGLEX-MWE Section's [standing committee]() (SC) consists of one elected officer (the MWE Section representative) and four nominated officers. The term of the SC nominated officer is 2 years, and term periods for different SC nominated officers should alternate, to ensure a transfer of competencies within the SC.

On July 7, 2024, a call for new standing committee officers was released on the SIGLEX-MWE members mailing list, and some potential candidates were also approached personally by the Section representative. As a result, three (3) expressions of interest were submitted via the online form by active members of the community who regularly publish research on multiword expressions.

After discussion within the standing committee, including both the former and newly elected Section representatives, a ranking was submitted to the SIGLEX board, who validated the nomination of Atul Ojha (University of Galway, Ireland) and Mathieu

Constant (Université de Lorraine, CNRS, ATILF, France) based on various balance factors including scientific excellence, diversity, seniority, linguistics/computer background, and multilingualism. The term of the two newly nominated members is from September 2024 to August 2026.

We thank the outgoing officers Kilian Evang (Heinrich Heine University Düsseldorf, Germany) and Lifeng Han (The University of Manchester, UK) for their valuable contributions. Both have joined the MWE Section's advisory committee.

Membership

As per the Section's constitution, SIGLEX-MWE members are those who have registered with SIGLEX and indicated their membership to the MWE Section in the registration form. As of December 31, 2024, the SIGLEX directory recorded 403 members registered under the MWE Section, compared to 365 members on December 31, 2023. These figures indicate the continued growth of the MWE Section. The publicly available SIGLEX directory may reflect a slightly lower number of members, as some members opt out of displaying their identity on the web.

MWE 2024 workshop

In line with the MWE Section's tradition of fostering synergies with other research communities, the 20th edition of the MWE workshop was co-organized by the ACL SIGLEX-MWE section and the Universal Dependencies (UD) community. The Joint Workshop on Multiword Expressions and Universal Dependencies (MWE-UD 2024) was held in Turin, Italy, on May 25, 2024, as a hybrid one-day event within the LREC-COLING Joint Conference.

The call for papers invited submissions of original research on MWEs, UDs, and their interplay. In particular, the following topics were mentioned as relevant: *sensitivity of LLMs to MWE and syntactic dependencies*; *applicability of UD and MWE annotation and discovery for low-resource and typologically diverse languages and language varieties*; *case studies on the consistency, coverage or universal applicability of MWE annotation in the UD or PARSEME frameworks, as well as studies on automatic detection and interpretation of MWEs in corpora; MWE and UD processing to enhance end-user applications including MT, simplification, language learning and assessment, social media mining, and abusive language detection; testing developed systems on the latest UD 2.13 and/or PARSEME 1.3 datasets.*

MWE-UD 2024 used START as the official submission platform; papers with reviews from OpenReview were also considered. A total of 54 submissions were received, 44 of which were archival, meaning original research papers (29 long and 15 short). Additionally, 10 papers were submitted as non-archival, that is papers presented elsewhere or reporting on work in progress. Non-archival submissions did not undergo further review and were not included in the workshop proceedings but were presented as posters.

Each of the archival submissions underwent review by three reviewers. A total of 26 archival submissions were selected for presentation. The review process was highly competitive, resulting in an acceptance rate of 59% for the archival submissions (as

opposed to 66.67% last year). One accepted paper was withdrawn.

The workshop's program included:

- 18 long papers (8 MWE, 6 MWE-UD, and 4 UD), 7 short papers
- Oral presentations: 15 (8 on-site, 7 online) and poster presentations (on-site): 20 posters;
- Tracks on MWEs, UDs, and a joint MWE-UD track;
- Two invited talks titled:
  - *Using Universal Dependencies for testing hypotheses about communicative efficiency* by [Natalia Levshina](#) (Radboud University, Nijmegen, The Netherlands), and
  - *Every Time We Hire an LLM, the Reasoning Performance of the Linguists Goes Up*, by [Harish Tayyar Madabushi](#) (University of Bath, UK).
- a community discussion, open to all.

As of 2023, the MWE workshop organizing committee decided to confer the Best Paper Award based on the reviews of accepted papers. The criteria for selecting the best paper are scientific significance, originality, technical quality, and clarity of presentation.

Participation in the workshop can only be roughly estimated since LREC-COLING 2024 organizers did not release the final numbers of registered participants. Based on the spreadsheet shared by the LREC-COLING 2024 organizers with all workshop organizers in early May, our workshop ranked 3rd overall in terms of registered participants and 2nd among one-day workshops with 79 registered participants. It was also ranked 1st for on-site registered participants (72 onsite and 7 online).

Proceedings were published in the [ACL Anthology](#).

Slides and posters are available on the website for most presentations, and we also share the [community discussion slides](#).

The MWE-UD 2024 workshop was co-organized and funded by SIGLEX and [UniDive COST action (CA21167)](#). Specifically, the SIGLEX board voted to provide SIGLEX events with $1500 each to spend towards increasing the diversity of attendees. 2 authors of accepted papers (out of five) who applied to receive subsidies from ACL towards onsite registration and travel costs were selected by the MWE-UD organizers. The selection was made based on geographic diversity, especially focusing on less attended regions, underrepresented groups (e.g., an underrepresented country), or low-resource languages.

UniDive COST Action provided funding for travel and daily costs for 29 participants, namely, the two invited speakers, workshop organizers who had not secured funding from elsewhere, plus authors of accepted papers—totaling 26 people out of the 27 who had applied. Unfortunately, due to COST regulations, only one author was not eligible and was not reimbursed.

MWE-related initiatives

In the past, members of the Section were involved in the organization of the [PARSEME](#)

shared tasks, co-located with the annual MWE workshop. The shared tasks were synchronized with the release of (training and evaluation) corpora annotated for verbal MWEs. In 2024, the Section did not organize any shared task, since the PARSEME network has worked towards disconnecting the corpus development, maintenance, and releases from the organization of shared tasks.

The PARSEME community has joined UniDive Cost Action CA21167, an interdisciplinary scientific network devoted to universality, diversity, and idiosyncrasy in language technology. Being highly idiosyncratic, MWEs lie at the heart of UniDive, along the axes of corpus annotation, representation in lexica (and the lexicon-corpus interface), and their efficient processing. Within UniDive, guidelines for annotating MWEs that belong to grammatical categories other than verbs (i.e., nominal, adjectival, adverbial and functional) are currently underway (PARSEME Guidelines version 2.0).

MWE 2025 workshop

A joint ACL/EACL/EMNLP call for workshop proposals was released on 9th September 2024 with the deadline set to 22nd October 2024. The proposal for the MWE 2025 workshop was submitted by the organizing committee members (Verginica Barbu Mititelu, Voula Giouli, A. Seza Dogruöz, Alexandre Rademaker, Atul Kr. Ojha, Gražina Korvel, Mathieu Constant). In our proposal, we had opted for either ACL2025 (Vienna, Austria) or NAACL2025 (Albuquerque, New Mexico, USA) – in this order – taking into account the results of a post-workshop survey sent after MWE 2021 (details in the 2021 yearly report), the availability of funds from UniDive only for events in Europe and also considering that EMNLP would be the least attractive venue for the targeted audience of our workshop.

MWE2025 was accepted as a hybrid one-day workshop at NAACL 2025.

The 1st call for papers was released in December 2024.