

Report on the activities of the SIGLEX-MWE Section in 2025

**Verginica Barbu Mititelu, A. Seza Dođruöz, Alexandre Rademaker, Mathieu Constant,
Atul Kr. Ojha, Ivelina Stoyanova, Adriana Pagano**

January 2026

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 2025.

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 June 29, 2025, 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 Ivelina Stoyanova (Institute for Bulgarian Language, Bulgarian Academy of Sciences) and Adriana Pagano (Universidade Federal de Minas Gerais, Brazil) 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 2025 to August 2027.

We thank the outgoing officers A. Seza Dođruöz (Ghent University, Belgium) and Alexandre Rademaker (IBM Research and FGV EMAP, Brazil) 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, 2025, the SIGLEX directory recorded 481 members registered under the MWE Section, compared to 403 members on December 31, 2024. These figures indicate the continued growth of the MWE Section.

MWE 2025 workshop

The 21st Workshop on Multiword Expressions (MWE 2025) was held in Albuquerque, New Mexico, U.S.A., on May 4, 2025, as a hybrid one-day event within the NAACL Conference.

The call for papers invited submissions of original research on MWEs. In particular, the following topics were mentioned as relevant: *MWE processing to enhance end-user applications; MWE processing and identification in the general language, as well as in specialized languages and domains; MWE processing in low-resource languages; MWE identification and interpretation in LLMs; new and enhanced representation of MWEs.*

MWE 2025 used OpenReview as the official submission platform. A total of 13 submissions were received, all archival, meaning original research papers. No non-archival papers were received.

Each of the archival submissions underwent review by three reviewers. A total of 9 archival submissions were selected for presentation. The acceptance rate was 69% for the archival submissions (as opposed to 59% last year). The workshop's program included:

- 5 long papers and 4 short papers;
- Oral presentations: 9 (1 on-site, 8 online);
- one invited talk titled Meaning Construction at the Syntax-Lexis Nexus by Nathan Schneider from Georgetown University
- a panel discussion on the topic of *Tokenization in the era of LLMs*. The panelists were: Agata Savary, (Université Paris-Saclay), Elizabeth Salesky (Google DeepMind), Jindřich Libovický (UFAL, Charles University), John P. McCrae (University of Galway), Yuval Pinter (Ben-Gurion University)
- 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. In 2025 the award went to Uliana Sentsova, Debora Ciminari, Josef van Genabith, Cristina España-Bonet for the paper *MultiCoPIE: A Multilingual Corpus of Potentially Idiomatic Expressions for Cross-lingual PIE Disambiguation*.

Participation in the workshop counted: 7 participants on-site and 17 online.

Proceedings were published in the ACL Anthology.

The MWE-UD 2025 workshop was co-organized and funded by SIGLEX with \$1500 each to spend towards increasing the diversity of attendees. 1 author of accepted papers and 3 students applied to receive subsidies from ACL towards online registration and were approved by the MWE 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.

MWE-related initiatives

AdMIRe 2025

With two subtasks: Static Images and Image Sequences - SemEval-2025 Shared Task on Advancing Multimodal Idiomaticity Representation

PARSEME 2.0

The MWE 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, in 2024 guidelines for annotating nominal, adjectival, adverbial and functional MWEs were created (PARSEME Guidelines version 2.0), while in 2025 a vast campaign of annotating them in corpora took place with an eye to organizing a new edition of the shared task on MWE identification in corpora. The evaluation took place in December 2025 and the results will be presented at MWE 2026 workshop.

AdMIRe 2.0

MWE 2026 Shared Task on Advancing Multimodal Idiomaticity Representation. It addresses the challenge of idiomatic language understanding by evaluating how well computational models interpret potentially idiomatic expressions (PIEs) using both text and images

MWE 2026 workshop

A joint ACL/EACL call for workshop proposals was released on 11th August 2025, with the deadline set to 5th September 2025. The proposal for the MWE 2026 workshop was submitted by the organizing committee members (Verginica Barbu Mititelu, A. Seza Dogruöz, Alexandre Rademaker, Atul Kr. Ojha, Mathieu Constant, Ivelina Stoyanova). In our proposal, we had opted for EACL 2026. MWE 2026 was accepted as a hybrid half-day workshop at EACL 2026. The 1st call for papers was released in October 2025.