

Report on the activities of the SIGLEX-MWE Section in 2022

Voula Giouli, Carlos Ramisch, Archna Bhatia, Paul Cook, Marcos Garcia, Shiva Taslimipoor, Kilian Evang, Lifeng Han

January 31, 2023

According to the constitution of the [SIGLEX-MWE Section](#), the standing committee of the Section should report yearly on the Section's activities to its members. This document is the report of the Section's activities in 2022.

SIGLEX elections – Section representative

Following the [SIGLEX constitution](#), the term of the SIGLEX board members elected in 2020, including the MWE section representative, finished in 2022, requiring the organisation of elections. A call for nominations was sent to all SIGLEX members on May 13, 2022 and was open until July 30, 2022. Preslav Nakov (SIGLEX president) and Carlos Ramisch (MWE Section representative) acted as returning officers, collected candidacy statements, and interacted with the ACL information officer. ACL provided the online elections' technical infrastructure on the *ElectionRunner* platform.

There was a single nomination for the position of the MWE Section representative. All 423 SIGLEX members registered until September 9, 2022 received an individual ballot email, and 146 SIGLEX members voted in the elections, which lasted from September 15 to October 15, 2022. Voula Giouli (ILSP Athens, Greece) was elected with 89% of the votes (11% abstained) thus becoming the MWE Section representative for the 2022-2024 term.

The other members of the standing committee 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 duration of the term of an SC nominated officer is 2 years, and term periods for different SC nominated officers should alternate, so as to ensure a transfer of competences within the SC.

On August 12, 2022, a call for new standing committee officers was announced on the MWE members mailing list, and some potential candidates were also approached personally by the Section representative. As a result, 6 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 Lifeng Han (University of Manchester, UK) and Kilian Evang (Heinrich Heine University Düsseldorf, Germany) on the basis of various balance factors including scientific excellence, diversity, seniority, linguistics/computer background and multilingualism. The term of the two newly nominated members is from October 2022 to September 2024.

We thank the outgoing officers Paul Cook (University of New Brunswick, Canada) and Shiva Taslimipoor (Cambridge University, UK) for their valuable contribution. Both have joined the MWE Section's [advisory committee](#).

Membership

According to the Section's constitution, SIGLEX-MWE members are those who have registered to SIGLEX and indicated their membership to the MWE Section in the registration form. On December 31, 2022 the [SIGLEX directory](#) indicated that 336 members registered as MWE Section members (290 members on Dec 31, 2021). These numbers indicate that the MWE Section keeps growing. The number of members might appear slightly lower on the public SIGLEX directory, as some members opt out of displaying their identity on the web.

MWE 2022 workshop

The [MWE 2022 workshop](#) took place as a hybrid event, with both online and on-site participants in Marseille, on June 25th, as part of LREC 2022. The workshop's program included:

- Oral and poster presentations of 6 short papers, 9 long papers, plus 5 presentations of non-archival papers (published elsewhere);
- 2 invited talks by Sabine Schulte im Walde and Steven Bird;
- a joint session with [SIGUL](#), the workshop of the ELRA/ISCA SIG for under-resourced languages, including a joint poster session and Steven Bird's invited talk;
- a community discussion, open to all.

The workshop venue was chosen taking into account the results of the post-workshop survey sent after MWE 2021 (details in the [2021 yearly report](#)). The goal was to favour networking, facilitating logistics in case of online or hybrid modes, being geographically close to a large part of the community (Europe), and being a conference with focus on resources, traditionally attended by a large part of the linguistics community.

The call for papers mentioned three main topics: *MWE processing in low-resource languages*, *MWE identification and interpretation in pre-trained language models*, and *MWE processing to enhance end-user applications*. The latter was also emphasised in 2021, being reconducted this year given its importance to connect the field to other NLP applications. The focus on low-resource languages was materialised by a joint session with the [2022 workshop of the Special Interest Group in Under-resourced Languages \(SIGUL\)](#).

The workshop had an attendance of around 20 participants online and around 20 participants onsite. Slides and posters are available on the website for most presentations, and we also

share the [community discussion](#) slides. Proceedings were published in the [ACL Anthology](#).

MWE-related initiatives

In 2022, the Section did not organise any shared task. In the past, members of the Section were involved in the organisation of the [PARSEME shared tasks](#), co-located with the annual MWE workshop. The shared tasks were synchronised with the release of (training and evaluation) corpora annotated for verbal MWEs. The [PARSEME network](#) is currently working towards disconnecting the corpus development, maintenance and releases from the organisation of shared tasks. The next corpus release is planned for the first semester of 2023.

The standing committee endorsed and gave feedback on the organisation of [SemEval 2022 task 2](#), on Multilingual Idiomaticity Detection and Sentence Embedding.

MWE 2023 workshop

A joint ACL/EACL/EMNLP call for workshop proposals was released late October 2022 with a (tight) deadline set to November 15th, 2022. The proposal for the MWE 2023 workshop was submitted by the organising committee members (Voula Giouli, Archana Bhatia, Kilian Evang, Marcos Garcia, Lifeng Han, and Shiva Taslimipoor). For this 19th edition of the workshop, we seek synergies with the Clinical NLP community; therefore a joint track/session was foreseen in the proposal.

The [MWE 2023](#) workshop was accepted as a hybrid one-day workshop at EACL 2023, in Dubrovnik, Croatia (May 5 or 6, 2023). The 1st call for papers was released in December 2022. Since the Clinical NLP workshop was accepted at ACL (Toronto, Canada), we plan to organise a joint track instead of a session; we also plan to invite a member of the Clinical NLP community to give a keynote speech at the MWE 2023 workshop.