

# Report on the Activities of the SIGLEX-MWE Section in 2018

*Agata Savary, Francis Bond, Stella Markantonatou, Jelena Mitrović and Carla Parra Escartín*

*25 February 2019*

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 2018.

## Standing Committee

Since August 2018, the SIGLEX-MWE Section has its [Standing Committee](#), consisting of one elected and four nominated officers. According to the [Constitution](#), 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. Since 2017-2018 was the first year of existence of the SC, two SC members were to be replaced already after a one-year term. In August 2018, a call for new SC candidates was announced, and 2 expressions of interest were submitted. Those were validated by the SIGLEX Board in September 2018. As a result, the two Officers stepping down were Carlos Ramisch and Johanna Monti, while the two newly nominated Officers are Jelena Mitrović and Carla Parra Escartín.

## Membership

The Section faces some difficulties in maintaining a precise list of its members. According to the Constitution, the SIGLEX-MWE members are those who have registered to SIGLEX with the "MWE" status. This list now contains 278 registered identities, some of which are obsolete or duplicated. On the other hand, the Section maintains a mailing list for its announcements ([multiword-expressions@lists.sourceforge.net](mailto:multiword-expressions@lists.sourceforge.net)) whose members do not fully overlap with the SIGLEX-MWE members. Additionally, membership management on this list is no longer possible, since on 10 June 2016, for the reasons of privacy protection, Sourceforge no longer allows the list administrators to view or change the names and addresses of their members.

Despite these obstacles, some cleaning-up actions were undertaken in spring 2018 to homogenise the SIGLEX-MWE and the [multiword-expressions@lists.sourceforge.net](mailto:multiword-expressions@lists.sourceforge.net) membership. A more reliable solution should be proposed in 2019.

## LAW-MWE-CxG 2018 workshop

The Section joined forces with the Special Interest Group for Annotation ([SIGANN](#)) to organise a joint even, the Joint Workshop on Linguistic Annotation, Multiword Expressions and

Constructions ([LAW-MWE-CxG-2018](#)). This joint event was co-located with the COLING 2018 conference in Santa Fe (USA) on August 25-26, 2018. It brought together the 12th Linguistic Annotation Workshop (LAW XII) and the 14th Workshop on Multiword Expressions (MWE).

- Participation: over 60 participants
- 3 invited talks:
  - "Annotation Schemes for Surface Construction Labeling" by Lori Levin, Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, USA.
  - "From Lexical Functional Grammar to Enhanced Universal Dependencies", by Adam Przepiórkowski, Institute of Philosophy, University of Warsaw, and Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland
  - "Leaving no token behind: comprehensive (and delicious) annotation of MWEs and supersenses" by Nathan Schneider, Georgetown University, USA
- Submissions: 34 papers (22 long and 12 short) were submitted to the research track; 16 long papers and 6 short ones were selected (9 as oral presentations, and 13 as posters), and 1 paper was withdrawn. The overall selectivity rate was 65%. 8 system description papers were submitted to the shared task track. All were selected.
- Proceedings: published in the [ACL Anthology](#).
- Shared task: the workshop was the culminating event of edition 1.1 of the PARSEME [shared task](#) on automatic identification of verbal MWEs (see below). The shared task organization and results were presented orally, and 8 out of the 12 participating systems were presented as posters.
- Panel discussion (see [slides](#)): it included feedback from this joint event, MWE-related announcements, announcements about the SIGLEX-MWE Steering Committee recruitment (see above), plans for future shared task editions, for UD-PARSEME collaboration, as well as for a centralized documentation of research questions of the MWE community.

## PARSEME shared task on automatic identification of verbal MWEs - edition 1.1

Edition 1.1 of this shared task was a follow-up of edition 1.0, whose culminating event took place in the MWE 2017 workshop at EACL in Valencia. Edition 1.1 used a new version of the PARSEME multilingual corpus of verbal multiword expressions, with an extended number of languages, enhanced annotation guidelines, and newly annotated texts. This initiative started in the framework of the IC1207 [PARSEME](#) COST action, and since the end of action (in April 2017) was overtaken by the SIGLEX-MWE section.

[Edition 1.1](#), thanks to a considerable effort of 20 language teams, resulted in:

- Universal [guidelines](#), enhanced with respect to edition 1.0, with examples in many languages and room for language-specific specifications.
- The final [corpus](#) of over **6 million tokens** and over **79,000 VMWE annotations in 20**

**languages**, distributed under different versions of the Creative Commons license (except for Arabic, covered by the LDC license) via the CLARIN/LINDAT infrastructure.

- 17 systems (from 12 teams) for VMWE identification (most of them were multilingual, 19 languages were covered) - see the [results](#)
- Novel evaluation measures (dedicated to specific phenomena identified as challenging in MWE processing)
- [cupt](#) file format (extending the conllu format), validators, converters and evaluation tools
- Data repositories (gitlab)

## PMWE volume with selected papers from the MWE 2017 workshop

A volume containing extended and reviewed versions of 13 selected papers from the MWE 2017 workshop was published in October 2018 in the "[Phraseology and Multiword Expressions](#)" book series at [Language Science Press](#):

- Stella Markantonatou, Carlos Ramisch, Agata Savary and Veronika Vincze (Eds.) "[Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop](#)", Language Science Press, Berlin, 2018.

## MWE-WN 2019 workshop proposal

In order to achieve better synergies with related initiatives, the Section made contact with the Global Wordnet Association ([GWA](#)). A joint workshop proposal, entitled "Joint Workshop on Multiword Expressions and WordNet ([MWE-WN 2019](#))", was submitted to ACL/COLING/EMNLP/NAACL 2019 and selected to be colocated with [ACL 2019](#) in Florence, Italy on August 2nd, 2019. The nomination of the workshop organisation and PC chairs was handled jointly by the SIGLEX-MWE Standing Committee and the GWA leaders. The Program Committee gathers experts from both communities.