

Introduction

Franciska de Jong

Executive Director CLARIN ERIC
Universiteit Utrecht, The Netherlands
f.m.g.dejong@uu.nl

Kiril Simov

IICT, Bulgarian Academy of Sciences
Sofia, Bulgaria
Programme Committee Chair
kivs@bultreebank.org

This volume presents the highlights of the 8th CLARIN Annual Conference 2019 held in Leipzig, Germany, on 30th September —2nd October 2019.

CLARIN ERIC¹ is the European Research Infrastructure for Language Resources and Technology aimed at supporting researchers from the Social Sciences and Humanities (SSH) and beyond in their use of language data and technologies. CLARIN works towards lowering barriers in doing research by giving access to language resources distributed across the countries involved in the infrastructure and by offering advanced, user-friendly and effective applications that enable the analysis of textual data, speech recordings, as well as multimodal material in a wide diversity of research tasks.

Since the establishment of the ERIC in 2012, CLARIN has grown in size considerably. Currently there are 21 member countries, 3 observers, and more than 100 associated research institutions who are all encouraged and supported to be represented at the annual conference which is meant to be a central event for CLARIN community and which is one of the crucial instrument for CLARIN to function as a knowledge hub. At the conference, consortia from all participating countries and the various communities of use meet, in order to exchange ideas, experiences and best practices in using the CLARIN infrastructure. The conference covers a wide range of topics, including the design, construction and operation of the CLARIN infrastructure, the data, tools and services that are or could be on offer, its actual use by researchers, its relation to other infrastructures and projects, and the CLARIN Knowledge Infrastructure. The aim is to attract researchers from all the various SSH fields that work with language materials, i.e. the people who are the *raison d'être* for CLARIN.

For the 8th edition of the CLARIN Annual Conference the special topic was “Humanities and Social Science research enabled by language resources and technology”. Early in 2019 a call² was issued for which 56 abstracts were submitted.

All submissions were reviewed anonymously by three reviewers (PC members and reviewers invited by PC members). Out of the 56 submitted abstracts 44 submissions were accepted for presentation at the conference (acceptance rate 0.79). The three topics that attracted the most of proposals were (a) language resources and tools, (b) design and construction of the CLARIN infrastructure, and (c) Interoperability and technical issues. This year 14 papers with a link to the special topic were accepted. Three of them had the special topic as the main focus, and thus formed the core of the plenary session devoted to it, while the others were presented either in other oral sessions or as posters. The accepted contributions were published in the online Proceedings of the Conference³. Several papers reported collaborative efforts by researchers from different institutes, either from one country or from institutional nodes in multiple countries. Some of the papers describe work that was carried out together with researchers from countries outside of Europe, such as South Africa, Japan, and Russia.

Following the well received novelty introduced at the 2018 edition of the CLARIN Annual Conference, a student poster session was organised with 13 presentations by PhD students who were selected by the national coordinator of their country. The abstracts of the student presentations were published in the online CLARIN 2019 Book of Abstracts⁴.

¹<http://clarin.eu>

²<https://www.clarin.eu/content/call-abstracts-clarin-annual-conference-2019>

³<https://office.clarin.eu/v/CE-2019-1512.CLARIN2019.ConferenceProceedings.pdf>

⁴<https://www.clarin.eu/clarin-annual-conference-2019-abstracts>

The conference hosted two invited talks:

- Professor Scott Rettberg (University of Bergen, Norway) introduced the audience to the various aspects of Electronic Literature: Documenting and Archiving Multimodal Computational Writing. In his presentation he addressed efforts to disseminate, document, and archive the field of electronic literature. After providing some examples of genres of electronic literature, and a number of projects was presented such as the Electronic Literature Collections, the ELMCIP Electronic Literature Knowledge Base, and the Electronic Literature Archive that seek to preserve a corpus of work and criticism for the future.
- Professor Elke Teich (University of the Saarland, Saarbrücken, Germany) gave a talk dedicated to corpus-driven investigation of language use, variation and change. Taking the perspective of a “humanist-as-scientist”, in her talk she reflected on the requirements for empirically investigating language use, variation and change with special regard to computational resources, models and tools.

In April 2019 CLARIN has launched the CLARIN Ambassadors Programme⁵ with the aim to raise awareness about CLARIN in disciplines and communities that are not yet fully familiar with the potential benefits of using the CLARIN infrastructure. One of the three appointed Ambassadors, Toine Pieters (Utrecht University, The Netherlands) gave an invited talk at the conference that was entitled “Towards a Universe of Local Time Machines - building an open eco-system for applied heritage fuelled by common language resources and existing infrastructure”.

In addition, on the event page⁶ CLARIN published a rich set of materials related to the conference:

- The complete conference programme and most of the slides presented: <https://www.clarin.eu/content/programme-clarin-annual-conference-2019>
- Recordings of most talks. (Note that the two invited lectures and several other video materials are available on a dedicated channel of VideoLectures: http://videolectures.net/clarinannualconference2019_leipzig/)

After the conference, the authors of the accepted papers and student submissions were invited to submit full versions of their papers to be considered for the post-conference proceedings volume. The papers were anonymously reviewed, each by three PC members. We received 18 (including 1 student paper) full length submissions, out of which 16 were accepted for this volume. All the main topics addressed at the conference are covered in this volume.

We would like to thank all PC members and reviewers for their efforts in evaluating and re-evaluating the submissions, Maria Eskevich from CLARIN Office for her indispensable support in the process of preparing these proceedings, and at Linköping University Electronic Press, who has ensured that the digital publication of this volume came about smoothly.

Members of the Programme Committee for the CLARIN Annual Conference 2019:

- Lars Borin, Språkbanken, University of Gothenburg, Sweden
- António Branco, Universidade de Lisboa, Portugal
- Griet Depoorter, Institute for the Dutch Language, The Netherlands/Flanders
- Koenraad De Smedt, University of Bergen, Norway
- Roald Eiselen, South African Centre for Digital Language Resources, South Africa
- Tomaž Erjavec, Jožef Stefan Institute, Slovenia

⁵<https://www.clarin.eu/news/meet-clarin-ambassadors>

⁶<https://www.clarin.eu/event/2019/clarin-annual-conference-2019-leipzig-germany>

- Eva Hajičová, Charles University Prague, Czech Republic
- Erhard Hinrichs, University of Tübingen, Germany
- Nicolas Larrousse, Huma-Num, France
- Krister Lindén, University of Helsinki, Finland
- Monica Monachini, Institute of Computational Linguistics “A. Zampolli”, Italy
- Karlheinz Mörth, Institute for Corpus Linguistics and Text Technology, Austrian Academy of Sciences, Austria
- Costanza Navarretta, University of Copenhagen, Denmark
- Jan Odiijk, Utrecht University, The Netherlands
- Maciej Piasecki, Wrocław University of Science and Technology, Poland
- Stelios Piperidis, Institute for Language and Speech Processing (ILSP), Athena Research Center, Greece
- Eiríkur Rögnvaldsson, University of Iceland, Iceland
- Kiril Simov, IICT, Bulgarian Academy of Sciences, Bulgaria (Chair)
- Inguna Skadiņa, University of Latvia, Latvia
- Marko Tadić, University of Zagreb, Croatia
- Jurgita Vaičėnienė, Vytautas Magnus University, Lithuania
- Tamás Váradi, Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary
- Kadri Vider, University of Tartu, Estonia
- Martin Wynne, University of Oxford, United Kingdom

Additional reviewers of this volume:

- Laska Laskova, Sofia University “St. Kliment Ochridski”, Bulgaria