

SLTU-CCURL 2020
1st Joint SLTU (Spoken Language Technologies for Under-resourced languages)
and CCURL (Collaboration and Computing for Under-Resourced Languages)
Workshop

<http://sltu-ccurl-2020.ilc.cnr.it>

1st Call for Papers

Date

11-12 May 2020

to be held as part of the 12th edition of the Language Resources and Evaluation Conference (LREC 2020) at the Palais du Pharo in Marseille (France)

endorsed by SIGUL (<http://www.elra.info/en/sig/sigul>), ELRA and ISCA (to be confirmed)

Invited Speakers

- Alan Black (Carnegie Mellon University, USA)
- Teresa Lynn (ADAPT Centre, Dublin City University, Ireland)

Tutorials

SLTU-CCURL 2020 is pleased to offer two Tutorials, to be held at Université Aix-marseille, near the LREC 2020 venue, on 10 May 2020:

- **T1: Building ASR systems using the Kaldi toolkit** by Jan Trmal (John Hopkins University)
- **T2: Introduction to Handwritten Text Recognition with Transkribus** by Achim Rabus (University of Freiburg)

More details will be announced on the Workshop Web site.

Attendance to these Tutorials will be free of charge, but registration will be required for organisational purposes (since the number of attendees will be limited to 25 per Tutorial).

Workshop Description and Objectives

The 1st Joint SLTU and CCURL Workshop will be held on 11-12 May 2020 in Marseille (France) during LREC 2020.

Organized by SIGUL, a joint Special Interest Group of the European Language Resources Association (ELRA) and of the International Speech Communication Association (ISCA), this Joint Workshop will gather researchers working on speech processing and NLP for less-resourced languages.

We solicit papers and posters related to all areas of NLP, speech and computational linguistics, as well as those at the intersection with digital humanities and documentary linguistics, provided that they address less-resourced languages.

Example topics are the following ones:

- language resource development, acquisition and representation;
- linguistic theories, corpus development and resources;
- linguistic and cognitive studies;
- unsupervised discovery of linguistic units;
- code switched lexical modeling;
- multi-lingual and cross-lingual (spoken, text) language processing;
- speech-to-text, text-to-speech and speech-to-speech processing;
- machine translation and dialogue systems;
- NLP and speech technologies for under-resourced languages.

The intention of this Joint Workshop is not only to provide a forum for the presentation of research, but also to offer a venue where researchers in different disciplines and from varied backgrounds can fruitfully explore new areas of intellectual and practical development while honoring their common interest of

sustaining less-resourced languages.

We will have both oral presentation sessions and poster sessions. Authors will be requested to specify the desired presentation mode at submission time. The decision on whether a presentation will be an oral or poster one will be taken by the Organizing Committee on the advice of the Program Committee, taking into account the subject matter and how it might be best conveyed. Oral and poster presentations will not be distinguished in the Workshop Proceedings.

Submission and Publication

Papers need to address less-resourced languages. They can contain an analysis and insight into existing methods and problems; a description of resources; an overview of the literature or of the current initiatives, or a combination of the above. Authors must declare if part of the paper contains material previously published elsewhere.

Prospective authors are invited to submit full-length papers of 8 pages (references excluded), strictly complying with the LREC stylesheet (<https://lrec2020.lrec-conf.org/en/submission2020/authors-kit>).

Papers should be submitted in PDF unprotected format to the Workshop START page (URL will be provided in due time).

Each submission will be reviewed by three Programme Committee members. In compliance with the LREC rules, papers must not be anonymized.

Accepted papers will be presented either as oral presentations or posters and will be published in the Workshop Proceedings.

The formatting template must be strictly adhered to and deadlines met.

Important Dates

- 14 February 2020: Paper submission deadline
- 13 March 2020: Paper notification of acceptance
- 2 April 2020: Camera-ready papers due
- 11-12 May 2020: Workshop date

Identify, Describe and Share your LRs!

Describing your LRs in the LRE Map is now a normal practice in the submission procedure of LREC (introduced in 2010 and adopted by other conferences). To continue the efforts initiated at LREC 2014 about "Sharing LRs" (data, tools, web- services, etc.), authors will have the possibility, when submitting a paper, to upload LRs in a special LREC repository. This effort of sharing LRs, linked to the LRE Map for their description, may become a new "regular" feature for conferences in our field, thus contributing to creating a common repository where everyone can deposit and share data.

As scientific work requires accurate citations of referenced work so as to allow the community to understand the whole context and also to allow the replication of the experiments conducted by other researchers, LREC 2020 endorses the need to uniquely Identify LRs through the use of the International Standard Language Resource Number (ISLRN, <http://www.islrn.org>), a Persistent Unique Identifier to be assigned to each Language Resource. The assignment of ISLRNs to LRs cited in LREC papers will be offered at submission time.

Workshop Chairs

- Dorothee Beermann (NTNU, Norway)
- Laurent Besacier (LIG – Université Grenoble Alpes, France)
- Sakriani Sakti (NAIST, Japan)
- Claudia Soria (CNR-ILC, Italy)

Programme Committee

- Adrian Doyle (Galway University, Ireland) TBC
- Alexey Karpov (SPIIRAS, Russian Federation)
- Alexis Palmer (University of North Texas, USA)

- Amita Dev (IGDTUW, India)
- Amir Aharoni (Wikimedia Foundation)
- Andras Kornai (Hungarian Academy of Sciences, Hungary)
- Angelo Mario Del Grosso (CNR-ILC, Italy)
- Antti Arppe (University of Alberta, Canada)
- Anupam Shukla (IIITM, India)
- Ayu Purwarianti (ITB, Indonesia) TBC
- Bruce Birch (The Minjilang Endangered Languages Publications Project, Australia) TBC
- Bruce Robertson (Mount Allison University, Canada) TBC
- Charl Van Heerden (SPbSU, Russian Federation) TBC
- Chiu Yu Tseng (ILAS, Taiwan) TBC Chris Cieri (LDC, USA) TBC
- Clara Rivera (Google) TBC
- Daan Van Esch (Google)
- Dafydd Gibbon (Bielefeld University, Germany)
- Delyth Prys (Bangor University, UK)
- Dewi Bryn Jones (Bangor University, UK)
- Dirk Van Compernelle (KU Leuven, Belgium) TBC
- Dorothee Beermann (NTNU, Norway)
- Emily Le Chen (University of Illinois, USA)
- Emily Prud'hommeaux (Boston College, USA) TBC
- Emmanuel Dupoux (EHESS-ENS, France) TBC
- Federico Boschetti (CNR-ILC, Italy)
- Francis Tyers (Moscow Higher School of Economics, Russia)
- Gerard Bailly (GIPSA Lab, CNRS) TBC
- Gilles Adda (LIMSI/IMMI CNRS, France)
- Hemant Patil (DA-IICT, India)
- Jeff Good (University at Buffalo, USA)
- John Judge (ADAPT DCU, Ireland)
- Jonas Fromseier Mortensen (Google)
- Jordan Lachler (University of Alberta, Canada) TBC
- Joseph Mariani (LIMSI-CNRS, France)
- Karunesh Arora (C-DAC, NOIDA, India) TBC
- Kepa Sarasola (University of the Basque Country, Spain)
- Kevin Scannell (Saint Louis University, Missouri, USA)
- Klara Ceberio (Elhuyar, Spain)
- Lane Schwartz (University of Illinois at Urbana-Champaign, USA)
- Laurent Besacier (LIG-IMAG, France)
- Lori Lamel (LIMSI, France) TBC
- Luong Chi-Mai (IOIT, Vietnam) TBC
- Maite Melero (Barcelona Supercomputing Center, Spain)
- Mans Hulden (University of Colorado Boulder, USA) TBC
- Miikka Silfverberg (University of Helsinki, Finland)
- Mikel Forcada (Universitat d'Alacant, Spain)
- Mirna Adriani (UI, Indonesia) TBC
- Mohammad A. M. Abushariah (The University of Jordan, Jordan)
- Nick Thieberger (University of Melbourne / ARC Centre of Excellence for the Dynamics of Language, Australia)
- Omar Farooq (AMU, India) Pierric Sans (Google) TBC
- Pradip K Das (IIT, India)
- Sakriani Sakti (NAIST, Japan)
- Satoshi Nakamura (NAIST, Japan)

- Sebastian Stüker (KIT, Germany)
- Shyam S Agrawal (KIIT, India)
- Sin Horng Chen (NCTU, Taiwan)
- Steven Bird (Charles Darwin University, Australia)
- Tan Tien Ping (USM, Malaysia) TBC
- Tanja Schultz (Uni-Bremen, Germany)
- Thang Vu (Uni-Stuttgart, Germany)
- Teresa Lynn (ADAPT Centre, Ireland)
- Trond Trosterud (Tromsø University, Norway)
- Tunde Adegbola (African Languages Technology Initiative, Nigeria)
- Vera Ferreira (CIDLeS - Interdisciplinary Centre for Social and Language Documentation, Portugal)
- Win Pa Pa (UCS Yangon, Myanmar)
- Xavier Anguera (Telefonica, Spain) TBC
- Yoshinori Sagisaka (Waseda University, Japan) TBC
- Zuraida Mohd Don (UPSI, Indonesia) TBC

Contact

- Laurent.Besacier@univ-grenoble-alpes.fr
- claudia.soria@ilc.cnr.it