

LREC 2020 – 12th International Conference on Language Resources and Evaluation
Opening Address
ELRA Secretary General and ELDA Chief Executive Officer

Khalid Choukri

Despite the distance and despite the current lockdown (almost everywhere), a great virtual welcome to everyone from the bottom of our hearts. We (the ELRA/ELDA team) hope that all our friends and their beloved ones are safe and in good spirit during these hard circumstances.

During the last few days and while writing this short message, my feelings went from the Provence atmosphere and Pagnol to Gabriel Garcia Marquez and love in times of Cholera; but with the sincere hope that our whole community (and beyond) is doing well.

22 years after the first LREC in Granada and for this first LREC in France, we had a very ambitious program: both to celebrate the 25th anniversary of ELRA and to share with all of you the best of Provence, one of the most beautiful French regions. Reviewing the last 25 years of ELRA activities serving our community was part of the plan.

This opportunity is lost for a while, LREC 2020 and the celebrations will not take place in May 2020.

It is with a very heavy heart that we had postponed LREC but we are confident that we will meet soon face to face with a chance to remember the best memories from previous LRECs.

Usually, the messages conveyed by the Committees take into account the nice atmosphere that we experience after a couple of days altogether to go beyond the written statements... It is difficult to carry the same sentiment in simple written words, but we all hope good days will come back soon and we, very likely and very hopefully, have learned more about caring for each other.

As you will read in all our messages, we do not give up. LREC is a major milestone in our lives and LREC is more than a simple scientific gathering, it is also one of the best moments when our community gets together, to share science but also friendship and memories.

So we are working hard to organize an LREC 2020 and we are confident that we will have the great pleasure to welcome you in Marseille and in person.

Let me share with you some of the key events and activities we carried out since we all met in Miyazaki.

LT4All in partnership with UNESCO

In December 2019, ELRA organized a 2-day event LT4All, Language Technologies for All, as part of the International Year of Indigenous Languages (IYIL2019), initiated by the United Nations. Recognizing that only 2% of the 7000 languages spoken in the world benefit from Language Technologies, the question asked was: what should be done for the remaining 98%, among which many are in danger of digital extinction, if not complete extinction.

With a strong commitment of ELRA, SIGUL, the ELRA-ISCA joint SIG on under-resourced languages, and Joseph Mariani, former ELRA President, now Honorary President, we managed to invite almost 400 participants from all continents and various backgrounds (language policy makers, language activists, and language technologists). Major players from our field, already aware of, and committed to these issues, took an active part in the conference. Over two days, challenges posed by the digital divide and multilingualism, particularly serious for indigenous languages, were reviewed. Then an overview of all the achievements that represent the state of the art in our field were presented to the large audience. Further to the success of the IYIL2019, including LT4All, followed by a plenary meeting

at the UN Headquarters in New York in 2019 and high-level UN meeting in Mexico in 2020, with ELRA contribution to both of which, the United Nations have decided to dedicate the next decade (2020-2032) to indigenous languages. These developments are very important for ELRA that gained visibility from these events, acknowledging its long-time involvement in providing and promoting Language Technology support for all languages.

As part of this involvement, ELRA supports the organization of a dedicated workshop at LREC 2020 by its SIGUL, the 1st Joint SLTU (Spoken Language Technologies for Under-resourced languages) and CCURL (Collaboration and Computing for Under-Resourced Languages) Workshop.

ELRA, part of important European initiatives

ELRA is involved in several major European initiatives dealing with Language Resources and Language Technology.

ELG, **the European Language Grid**, is a very important project started in Europe, a major initiative that targets the deployment of a marketplace for Language Resources and Language Technology. ELRA will bring its experience in handling issues related to language resources sharing and distribution. We do expect ELRA members to have multiple benefits from the ELG platform as it will allow to efficiently access services and resources that helps develop new applications.

ELRA has continued to be involved in the series of ELRC (**European Language Resources Coordination**) activities, including the collection of language data for Machine Translation and the organization of reach-out workshops all over Europe (EC Member States, plus Iceland and Norway). In this framework and with a strong support from the European commission, the ELRC consortium has secured a large number of resources, in particular for Machine Translation activities. With the encouragement of the European Commission, ELRA has released 200+ resources usable immediately, freely and openly, by our community, offering another distribution channel to this important assets.

ELRA, part of an important initiative to help address issues related to COVID-19

ELRA, with a large number of partners, is also taking part in a COVID-19 initiative that aims at producing specific Languages Resources so the community can carry out research activities related to information retrieval, machine translation, summarization, fake news identification and checking, etc. Another dimension of this initiative is the setup of a number of challenges and evaluation campaigns to assess the performance of current technology with respect to all these applications and more. The ultimate goal being to ensure that our community contributes to safeguard the welfare of human beings using Language technologies.

Other ELRA activities

The activities around the Language Resources at ELRA encompass a number of aspects including the management of legal issues. At the moment, the number of Language Resources in the ELRA catalogue amounts to 1380. This includes large sets of Language Resources that ELRA is now distributing as part of agreements signed in 2019 and 2020 with international partners to offer more assets to our members and the community at large.

I would like to emphasize the large number of agreements we signed with data right holders to bring in more resources with a strong focus on free resources.

ELRA also continues to assist or partners and members in all legal analysis and negotiations related to language resources to make sure that we are using resources that have been legally cleared and that don't

infringe any copyright or other right of third parties. We have also continued our activities in producing language resources in partnership with third parties. For instance we took part in the MGB challenge that required production of colloquial Arabic from North Africa. Our objective was to ensure that such resources were packaged and made available to the community at large.

Shared LRs at LREC 2018

The Languages Resources shared in 2018 have been made available from the LREC 2018 website (<https://bit.ly/3dpWdq3>). ELRA has long recognised the importance of sharing Language Resources (LRs) and making them available to the community. Since the 2014 edition of LREC, participants have been offered the possibility to share their LRs (data, tools, web-services, etc.) when submitting a paper, uploading them in a special LREC repository set up by ELRA. This effort of sharing LRs, linked to the LRE Map initiative (<https://lremap.elra.info>) for their description, contributes to creating a common repository where everyone can deposit and share data.

Acknowledgments

Last but not least I would like express my deep thanks to all our partners that continue to support the organization of LREC despite the current situation hoping like us for an upcoming face-to-face gathering. I would like to thank our Sponsors: Google, Amazon, Lenovo, Bertin, Ontotex, Vecsys and Vocapia.

I would like to thank Philippe Blache and Frédéric Béchet, the LREC Local Committee in Marseille for providing support to both the scientific content and the local organization during the preparatory phase for this LREC Edition in Marseille. Their efforts will not be wasted, our plans for an LREC 2020 in Marseille remain a top priority in our schedule for the months to come. Finally I would like to warmly thank the joint team of the two institutions that devoted so much effort over months and often behind curtains to make this one week memorable: ILC-CNR in Pisa and my own team, ELDA, in Paris. These are the two LREC coordinators and pillars: Sara Goggi and Hélène Mazo, and the team: Roberto Bartolini, Valérie Mapelli, Vincenzo Parrinelli, Caroline Rannaud, and Alexandre Sicard. I share their disappointment but their efforts will materialize when LREC takes place soon.

I look forward to the LREC 2020 face to face conference. In the meantime, I hope that you will enjoy the fruitful workshops that are going virtual, and ELRA is happy to assist in this. I also hope that these proceedings will reflect the creativity of our community.

I hope that each of you is safe, and so is your family and friends.

Paris, France, May 13, 2020

Khalid Choukri

ELRA Secretary General and ELDA Chief Executive Officer