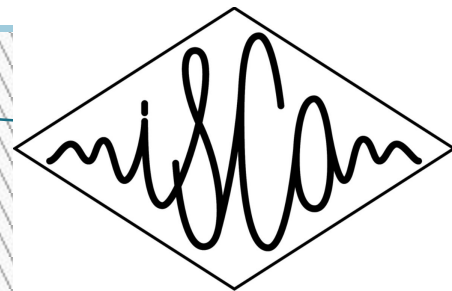
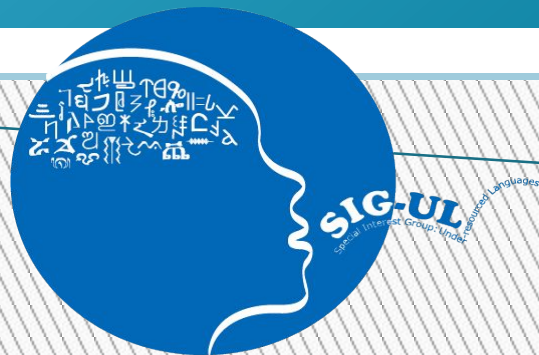


JOINT ISCA & ELRA

Special Interest Group Under-resourced Languages (SIGUL)

<http://www.elra.info/en/sig/sigul/>



SIG UNDER-RESOURCED LANGUAGES (SIGUL)

■ Joint Special Interest Group

✓ ISCA & ELRA

■ Main objectives

✓ Supports linguistic diversity through technology and ICT

✓ Commits to increasing the lesser-resourced languages (regional, minority, or endangered) chances to survive the digital world

■ Board (after elections end of 2017)

✓ Co-chair and ELRA liaison representative: Claudia Soria (CNR-ILC, Pisa, Italy)

✓ Co-chair and ISCA liaison representative: Laurent Besacier (LIG, France)

✓ Secretary: Sakriani Sakti (NAIST, Nara, Japan)

✓ Contributor: Dorothee Beermann (NTNU, Norway) + committee and international advisory group



SIGUL 2017 Activities

■ First Event: Interspeech 2017 Special Session

“Digital Revolution for Under-resourced Languages (DigRevURL)”

- ✓ One full day Interspeech 2017 Special Session (23 August 2017, Stockholm, Sweden)
- ✓ 3 Sessions: oral, poster, and panel discussion
- ✓ Presentation: 22 regular papers (6 oral, 16 poster) + 9 invited poster from the panelist
- ✓ Panel discussion (Co-chairs: Joseph Mariani and Sakriani Sakti) with 13 Panelists from ISCA/ELRA SIGUL, SLTU Board, Zero Resource Challenge, BABEL, LORELEI, BULB, JHU Summer Workshops, Google Research, UNESCO, and 3 region representative (Europe, African, Asian)



SIGUL 2018 Activities (1/3)

■ Start over the SIG recently created

- ✓ Mailing list: sigul@list.elra.info (>200 members)
- ✓ Sign up on: <https://bit.ly/2PYhM66>
- ✓ Web page: <http://www.elra.info/en/sig/sigul>
- ✓ International advisory group

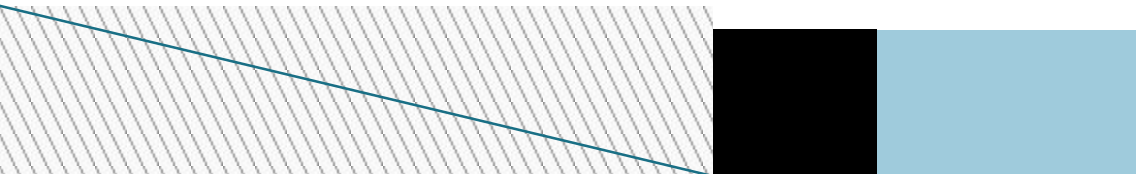
■ LREC 2018 Satellite Workshop (May 2018)

- ✓ “Collaboration and Computing for Under-Resourced Languages Towards an Alliance for Digital Language Diversity (CCURL)”
- ✓ Miyazaki, Japan
- ✓ 1 keynote, 15 papers, 30 participants – proceedings now online
<http://lrec-conf.org/workshops/lrec2018/W26/>

SIGUL 2018 Activities (2/3)

■ INTERSPEECH 2018 Satellite Workshop (September 2018):

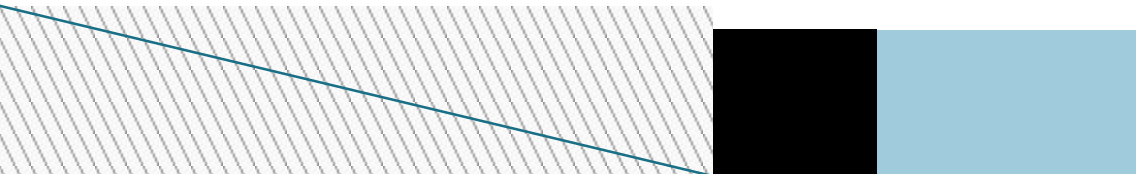
- ✓ “Workshop on Spoken Language Technologies for Under-Resourced Languages (SLTU)”
- ✓ New Dehli, India
- ✓ 72 papers submitted (after removing out-of-domain) ; half from India+SAARC ; 54 papers accepted
- ✓ 2 tutorials organized (30 participants each): (1) ASR using Kaldi toolkit (2) Neural Machine Translation
- ✓ 2 keynotes: (1) Pushpak Bhattacharyya (machine translation) (2) Emmanuel Dupoux (zero resource speech research)
- ✓ 1 panel discussion with industrial and academic participants, focus on spoken language technologies in India and SAARC countries



SIGUL 2018 Activities (3/3)

■ Participate in UNESCO Events

- ✓ **International Conference “Role of linguistic diversity in building a global community with shared future: protection, access and promotion of language resources”**
- ✓ **19 – 21 September 2018, Changsha, The People’s Republic of China**
- ✓ **Presentation:** Introduced SIGUL and language technologies for under-resourced languages
- ✓ **Involved in drafting outcome document:** "Protection and Promotion of Linguistic Diversity of the World" Yuelu Proclamation
- ✓ **Main point:** Combine the protection and promotion of linguistic diversity with the development of science and technology.



SIGUL 2019 Plans (1/2)

■ ICPHS 2019

- ✓ Special session accepted at ICPHS 2019 (Melbourne, Australia)
- ✓ Computational Approaches for Documenting and Analyzing Oral Languages
- ✓ Attracting computer scientists to ICPHS and engaging them in discussions with phoneticians (and linguists generally)
- ✓ <http://lig-getalp.imag.fr/icphs-2019-special-session/>

■ Local workshop in Uganda

- ✓ 2 researchers to Uganda to give a series of courses and organize data collection
- ✓ 1 in speech science and linguistics (Dorothee Beermann) ; 1 in speech technology (to be defined)
- ✓ 1 speaker may be funded by ISCA distinguished lecture series ; a second one by ELRA ?
- ✓ D. Beermann will prepare this through a self-funded trip to Uganda in Nov 2018

■ Zero Speech Challenge 2019

- ✓ SIGUL as a partner
- ✓ To be confirmed and set up

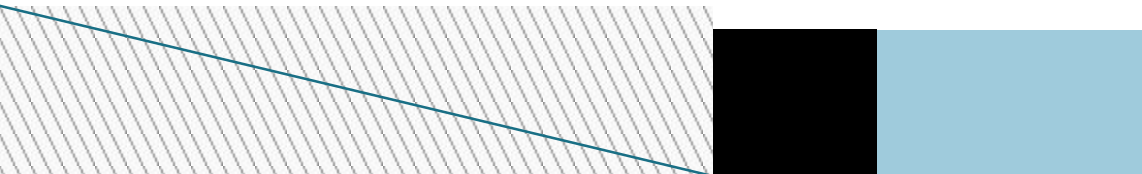
SIGUL 2019 Plans (2/2)

■ Local workshop in Japan

- ✓ "Digital Revolution for Under-resourced Languages in Asia Region", Feb 5–6, 2019, Nara, Japan
- ✓ **Objectives:** Activate under-resourced language communities in Asia
- ✓ **Invite 7 experts from Asia:** Indonesia, Malaysia, Phillipines, Thailand, Vietnam, India, China
- ✓ **Content:** Short presentation by each expert and panel discussion

■ Language Technology for All (LT4All)

- ✓ Related to 2019 UNESCO International Year of Indigenous languages
- ✓ **Participants based only on invitation** (academia, industries, governments, indigenous communities)
- ✓ **Content:** White papers on LT for UL + Several panel discussions
[Still in discussion with UNESCO – Not fixed yet]



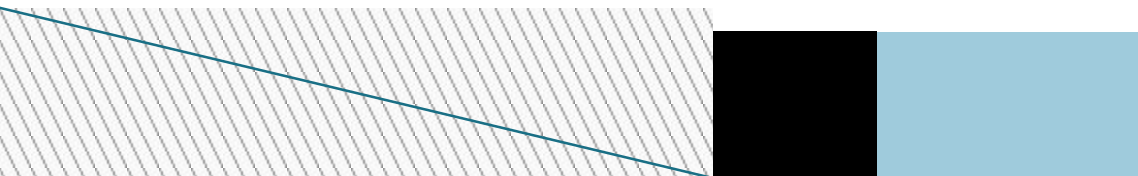
SIGUL 2020 Plans

■ Joint SLTU and CCURL Workshop in 2020 ?

- ✓ SLTU: spoken language community / CCURL: nlp community
- ✓ Should we merge ?
- ✓ Alternate satellite of Interspeech and LREC
- ✓ Joint workshop in France in 2020 as a satellite of LREC which will be in Marseille ?

■ Increase our visibility

- ✓ In South America, Africa, Eastern Europe, Central Asia ...
- ✓ First contacts with Brazil for an event (SLTU/CCURL?) in 2020



Budget Request 2019

Request to ELRA

✓ 5k€ for the following actions:

2k local workshop in Uganda (travel funding for 1 person)

2k local workshop in Japan (travel funding for 1 person)

1k for other actions (sigul meetings, zrc challenge)

