

LREC 2020 Workshop
Language Resources and Evaluation Conference
11–16 May 2020

**Workshop on Linguistic and Neurocognitive Resources
(LiNCR2020)**

PROCEEDINGS

Emmanuele Chersoni, Barry Devereux, Chu-Ren Huang. (eds.)

**Proceedings of the LREC 2020
Workshop on Linguistic and Neurocognitive Resources
(LiNCr2020)**

Editors: Emmanuele Chersoni, Barry Devereux, and Chu-Ren Huang.

ISBN: 979-10-95546-52-8

EAN: 9791095546528

For more information:

European Language Resources Association (ELRA)

9 rue des Cordelières

75013, Paris

France

<http://www.elra.info>

Email: lrec@elda.org

© European Language Resources Association (ELRA)

These Workshop Proceedings are licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License

Introduction

In the proceedings of the first LiNCR (pronounced ‘linker’) workshop in 2018, we stated that the aim of the workshop was to provide a venue to share and explore a new generation of language resources which link and aggregate cognitive, behavioural, neuroimaging and linguistic data. Our vision was to provide a forum for research that leads to the development of methods for the integration of neuro-cognitive data on language function with linguistic facts, the interpretation of experimental data when linked to rich linguistic information, and demonstrations of how new insights can be drawn from this powerful approach in domains such as language learning and neuro-cognitive deficits. We envisioned that there will be many future LiNCR workshops, just as the current 2nd workshop that we are presenting now.

What we did not foresee, however, was that we won’t be able to meet face-to-face and strengthen our links during this time of social distancing.

Nevertheless, the eight papers for presentation in this workshop will continue to showcase the innovative nature and potential for impact of this interdisciplinary and data-driven framework for understanding language and cognition. Three significant datasets are presented: the Little Prince corpus for neuro-cognitive studies in 26 languages, sensorimotor norms for Russian, and a dataset for complex emotion learning. Three papers focus on leveraging neuro-cognitive measurement for language technology, or vice versa. Finally, two papers deal with practical issues, such as fonts for dyslexia readers and language models for cloze task answer generation. The eclectic nature of these paper underlines both the vast frontiers to be explored yet as well as the versatility of linked linguistic and neuro-cognitive resources.

Most of all, we look forward to the third LiNCR to update our new findings, to overcome the temporary physical distances, and possibly even to show how linked linguistic and neuro-cognitive databases can shed light on issues related to epidemiology and public health.

Emmanuele Chersoni, Barry Devereux, and Chu-Ren Huang

May 2020

Organizers:

Emmanuele Chersoni (The Hong Kong Polytechnic University)
Barry Devereux (Queen's University Belfast)
Chu-Ren Huang (The Hong Kong Polytechnic University)

Program Committee:

Kathleen Ahrens (The Hong Kong Polytechnic University)
Amir Bakarov (Higher School of Economics, Moscow)
Leonor Becerra-Bonache (Jean Monnet University)
Philippe Blache (LPL-CNRS)
Zhenguang Cai (Chinese University of Hong Kong)
Cristopher Cieri (University of Pennsylvania)
Steven Derby (Queen's University Belfast)
Vesna Djokic (Goldsmiths University of London)
Stefan Frank (Radboud University of Nijmegen)
Diego Frassinelli (University of Stuttgart)
Francesca Frontini (Paul Valéry University of Montpellier)
John Hale (University of Georgia)
Nora Hollenstein (ETH Zurich)
Shu-Kai Hsieh (National Taiwan University)
Yu-Yin Hsu (The Hong Kong Polytechnic University)
Elisabetta Jezek (University of Pavia)
María Dolores Jiménez López (Universitat Rovira i Virgili, Tarragona)
Ping Li (The Hong Kong Polytechnic University)
Andreas Maria Liesenfeld (The Hong Kong Polytechnic University)
Yunfei Long (University of Nottingham)
Brian MacWhinney (Carnegie Mellon University)
Helen Meng (Chinese University of Hong Kong)
Karl David Neergaard (University of Macau)
Noel Nguyen (LPL-CNRS)
Mark Ormerod (Queen's University Belfast)
Christophe Pallier (INSERM-CEA)
Ludovica Pannitto (University of Trento)
Adam Pease (Articulate Software)
Massimo Poesio (Queen Mary University of London)
Stephen Politzer-Ahles (The Hong Kong Polytechnic University)
James Pustejovsky (Brandeis University)
Giulia Rambelli (University of Pisa)
Anna Rogers (University of Massachusetts Lowell)
Marco Senaldi (Scuola Normale Superiore, Pisa)
Adrià Torrens Urrutia (Universitat Rovira i Virgili, Tarragona)
Marten Van Schijndel (Cornell University)
Cory Shain (Ohio State University)
Mingyu Wan (The Hong Kong Polytechnic University)

Table of Contents

<i>Extrapolating Binder Style Word Embeddings to New Words</i> Jacob Turton, David Vinson and Robert Smith	1
<i>Towards the First Dyslexic Font in Russian</i> Svetlana Alexeeva, Aleksandra Dobrego and Vladislav Zubov	9
<i>Towards Best Practices for Leveraging Human Language Processing Signals for Natural Language Processing</i> Nora Hollenstein, Maria Barrett and Lisa Beinborn	15
<i>Language Models for Cloze Task Answer Generation in Russian</i> Anastasia Nikiforova, Sergey Pletenev, Daria Sinitsyna, Semen Sorokin, Anastasia Lopukhina and Nick Howell	28
<i>Does History Matter? Using Narrative Context to Predict the Trajectory of Sentence Sentiment</i> Liam Watson, Anna Jurek-Loughrey, Barry Devereux and Brian Murphy	38
<i>The Little Prince in 26 Languages: Towards a Multilingual Neuro-Cognitive Corpus</i> Sabrina Stehwien, Lena Henke, John Hale, Jonathan Brennan and Lars Meyer	43
<i>Towards a Multi-Dataset for Complex Emotions Learning Based on Deep Neural Networks</i> Billal Belainine, Fatiha Sadat, Mounir Boukadoum and Hakim Lounis	50
<i>Sensorimotor Norms for 506 Russian Nouns</i> Alex Miklashevsky	59