

EACL 2021

**The 16th Conference of the European Chapter of
the Association for Computational Linguistics**

Proceedings of the Student Research Workshop

April 19 - 23, 2021

©2021 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-954085-04-6

Introduction

Welcome to the EACL 2021 Student Research Workshop!

The EACL 2021 Student Research Workshop (SRW) is a platform for students in the field of Computational Linguistics and Natural Language Processing to come together to discuss and advance their research with help from more experienced researchers from both academia and industry.

The workshop is uniquely placed to provide valuable feedback to students about their research both before and after paper submissions. It provides them with ample opportunities to improve writing and research dissemination skills in the process. Following the tradition of the previous student research workshops, we have two tracks: research papers and thesis proposals. The research paper track is a venue for Ph.D. students, Master's students, and advanced undergraduate students to describe completed work or work-in-progress along with preliminary results. The thesis proposal track is offered for advanced Masters and Ph.D. students who have decided on a thesis topic and are interested in receiving feedback for their proposal with suggestions for both making the ideas achievable, as well as discussions related to future directions for their work.

The student research workshop has received considerable attention from many different parts of the world, and papers have addressed research questions in several different languages, which reflects the growth of the workshop. After excluding the withdrawn submissions, we received 59 submissions in total: 6 thesis proposals and 53 research papers. We accepted 4 thesis proposals and 22 research papers, resulting in an overall acceptance rate of 44%. We have also added a best paper award in the process. All the accepted papers will be presented virtually, as a part of the EACL conference, spread across three days (April 21st–23rd, 2021).

Mentoring is a core part of the SRW. In keeping with previous years, we organized pre-submission mentoring to improve the writing style and presentation of submissions. A total of 11 papers participated in this program. It offered students the opportunity to receive guidance from an experienced researcher before their submission was peer-reviewed for acceptance.

We thank our program committee members for providing careful and comprehensive reviews for the papers, and all of our mentors for donating their time to provide feedback to our student authors. Thanks to our faculty advisor, Eneko Agirre, for the essential advice and suggestions, and to the EACL 2021 organizing committee for their support in the entire process. Finally, we would like to thank all the authors whose participation has made the workshop a success!

Organizers:

Ionut-Teodor Sorodoc, Pompeu Fabra University

Madhumita Sushil, University of Antwerp

Ece Takmaz, University of Amsterdam

Faculty advisor:

Eneko Agirre, University of the Basque Country

Pre-submission Mentors:

Raquel Fernandez, University of Amsterdam

German Kruszewski, Naver Labs Europe

Xiaoyu Shen, Amazon

Valerio Basile, University of Turin

Bhuwan Dhingra, Google

Jeska Buhmann, University of Antwerp

Eva Vanmassenhove, Tilburg University

Alan Akbib, Humboldt-Universität zu Berlin

Hamed Zamani, University of Massachusetts Amherst

Sudipta Kar, Amazon Alexa AI

Lifu Huang, Virginia Tech

Fernando Alva Machego, University of Sheffield

Chris Brew, LivePerson

Program Committee:

Giovanni Cassani, Tilburg University

Sabine Schulte im Walde, University of Stuttgart

Thomas Brochhagen, Universitat Pompeu Fabra

Odette Scharenborg, Delft University of Technology

Sebastian Schuster, Stanford University

Sepideh Sadeghi, Google

Sebastian Pado, University of Stuttgart

Iacer Calixto, New York University

Vera Demberg, Saarland University
Ehsan Kamalloo, University of Alberta
Chris Brew, LivePerson
Lina M. Rojas-Barahona, Orange-Labs
Thomas Kleinbauer, Saarland University
Chaitanya Shivade, Amazon
Anusha Balakrishnan, Microsoft Semantic Machines
Thomas Kober, Rasa
Bonnie Webber, University of Edinburgh
Vered Shwartz, AI2 & University of Washington
Emily Bender, University of Washington
Andrea Horbach, University Duisburg-Essen
De Clercq Orphée, Ghent University
Sowmya Vajjala, National Research Council, Canada
Bruno Martins, University of Lisbon
Roberto Basili, University of Roma
Robert Logan, University of California, Irvine
Jacob Eisenstein, Google
Sewon Min, University of Washington
Nitish Gupta, University of Pennsylvania
Oshin Agarwal, University of Pennsylvania
David Adelani, Saarland University
Jack Hessel, AI2
Marius Mosbach, Saarland University
Enrique Manjavacas, University of Antwerp
Folger Karsdorp, KNAW Meertens Instituut
Florian Kunneman, Vrije Universiteit Amsterdam
Mona Jalal, Boston University
Lucy Lin, University of Washington
Paul Rayson, Lancaster University
Shruti Rijhwani, Carnegie Mellon University
Arnab Bhattacharya, IIT Kanpur

Nivedita Yadav, Janssen Pharmaceutica
Timothy Baldwin, The University of Melbourne
Esther van den Berg, Heidelberg University
Lucas Weber, University Pompeu Fabra
Tim Kreutz, University of Antwerp
Carolina Scarton, University of Sheffield
Fernando Alva-Manchego, University of Sheffield
Sian Gooding, University of Sheffield
Sebastian Gehrmann, Google Research
Albert Gatt, University of Malta
Ben Zhou, University of Pennsylvania
Kevin Lin, Microsoft
Denis Newman-Griffis, University of Pittsburgh
Omid Memarrast, University of Illinois of Chicago
Marta Ruiz, Universitat Politècnica de Catalunya
Alina Karakanta, Fondazione Bruno Kessler / University of Trento
A.B. Siddique, University of California, Riverside
Marco Damonte, Amazon
Alberto Testoni, University of Trento
Divyansh Kaushik, Carnegie Mellon University
Massimo Nicosia, Google
Bernhard Kratzwald, ETH Zurich
Florian Mai, Idiap Research Institute
Pieter Fizez, University of Antwerp
Chris Develder, Ghent University
Michael Hedderich, Saarland University
Louise Deleger, Université Paris-Saclay, INRAE
Filip Ilievski, USC Information Sciences Institute
Lifu Huang, Virginia Tech
Laura Ana Maria, University of Stuttgart
Alan Akbik, Humboldt-Universität zu Berlin
Najoung Kim, Johns Hopkins University

Vlad Niculae, University of Amsterdam
Diego Marcheggiani, Amazon
Michael Heddrich, Saarland University
Dirk Hovy, Bocconi University
Chris Alberti, Google
Roberto Dessi, Universitat Pompeu Fabra / Facebook AI Research
Louise McNally, Universitat Pompeu Fabra
Ru Meng, University of Pittsburgh
Phong Le, Amazon
Anna Currey, Amazon AWS AI
Arda Tezcan, Ghent University
Mattia Antonino Di gangi, AppTek
Katharina Kann , CU Boulder
Marcos Garcia, Universidade da Corunha
Murathan Kurfali, Stockholm University
Manex Agirrezabal, University of Copenhagen
Burcu Can, The University of Wolverhampton
Taraka Rama, University of North Texas
Ahmet Üstün, University of Groningen
Laura Aina, Pompeu Fabra University
Gosse Minnema, University of Groningen
Aina Gari, Université Paris-Saclay, CNRS, LIMSI
Ludovica Pannitto, University of Trento
Mario Giulianelli, University of Amsterdam
Qianchu Liu, University of Cambridge
Meishan Zhang, Tianjin University, China
Enrica Troiano, Institut für Maschinelle Sprachverarbeitung (IMS), Stuttgart
Shabnam Tafreshi, George Washington University
Zeerak Waseem, University of Sheffield
David Jurgens, University of Michigan
Els Lefever, Ghent University
Saadia Gabriel, University of Washington

Miguel A. Alonso, Universidade da Coruña

Hardy Hardy, The University of Sheffield

Bernd Bohnet, Google

Ruket Cakici, NTENT, METU

Table of Contents

<i>Computationally Efficient Wasserstein Loss for Structured Labels</i> Ayato Toyokuni, Sho Yokoi, Hisashi Kashima and Makoto Yamada	1
<i>Have Attention Heads in BERT Learned Constituency Grammar?</i> Ziyang Luo	8
<i>Do we read what we hear? Modeling orthographic influences on spoken word recognition</i> Nicole Macher, Badr M. Abdullah, Harm Brouwer and Dietrich Klakow	16
<i>PENELOPIE: Enabling Open Information Extraction for the Greek Language through Machine Translation</i> Dimitris Papadopoulos, Nikolaos Papadakis and Nikolaos Matsatsinis	23
<i>A Computational Analysis of Vagueness in Revisions of Instructional Texts</i> Alok Debnath and Michael Roth	30
<i>A reproduction of Apple’s bi-directional LSTM models for language identification in short strings</i> Mads Tofttrup, Søren Asger Sørensen, Manuel R. Ciosici and Ira Assent	36
<i>Automatically Cataloging Scholarly Articles using Library of Congress Subject Headings</i> Nazmul Kazi, Nathaniel Lane and Indika Kahanda	43
<i>Model Agnostic Answer Reranking System for Adversarial Question Answering</i> Sagnik Majumder, Chinmoy Samant and Greg Durrett	50
<i>BERT meets Cranfield: Uncovering the Properties of Full Ranking on Fully Labeled Data</i> Negin Ghasemi and Djoerd Hiemstra	58
<i>Siamese Neural Networks for Detecting Complementary Products</i> Marina Angelovska, Sina Sheikholeslami, Bas Dunn and Amir H. Payberah	65
<i>Contrasting distinct structured views to learn sentence embeddings</i> Antoine Simoulin and Benoit Crabbé	71
<i>Discrete Reasoning Templates for Natural Language Understanding</i> Hadeel Al-Negheimish, Pranava Madhyastha and Alessandra Russo	80
<i>Multilingual Email Zoning</i> Bruno Jardim, Ricardo Rei and Mariana S. C. Almeida	88
<i>Familiar words but strange voices: Modelling the influence of speech variability on word recognition</i> Alexandra Mayn, Badr M. Abdullah and Dietrich Klakow	96
<i>Emoji-Based Transfer Learning for Sentiment Tasks</i> Susann Boy, Dana Ruitter and Dietrich Klakow	103
<i>A Little Pretraining Goes a Long Way: A Case Study on Dependency Parsing Task for Low-resource Morphologically Rich Languages</i> Jivnesh Sandhan, Amrith Krishna, Ashim Gupta, Laxmidhar Behera and Pawan Goyal	111
<i>Development of Conversational AI for Sleep Coaching Programme</i> Heereen Shim	121

<i>Relating Relations: Meta-Relation Extraction from Online Health Forum Posts</i> Daniel Stickley	129
<i>Towards Personalised and Document-level Machine Translation of Dialogue</i> Sebastian Vincent	137
<i>Semantic-aware transformation of short texts using word embeddings: An application in the Food Computing domain</i> Andrea Morales-Garzón, Juan Gómez-Romero and Maria J. Martin-Bautista	148
<i>TMR: Evaluating NER Recall on Tough Mentions</i> Jingxuan Tu and Constantine Lignos	155
<i>The Effectiveness of Morphology-aware Segmentation in Low-Resource Neural Machine Translation</i> Jonne Saleva and Constantine Lignos	164
<i>Making Use of Latent Space in Language GANs for Generating Diverse Text without Pre-training</i> Takeshi Kojima, Yusuke Iwasawa and Yutaka Matsuo	175
<i>Beyond the English Web: Zero-Shot Cross-Lingual and Lightweight Monolingual Classification of Registers</i> Liina Repo, Valterri Skantsi, Samuel Rönqvist, Saara Hellström, Miika Oinonen, Anna Salmela, Douglas Biber, Jesse Egbert, Sampo Pyysalo and Veronika Laippala	183
<i>Explaining and Improving BERT Performance on Lexical Semantic Change Detection</i> Severin Laicher, Sinan Kurtiyigit, Dominik Schlechtweg, Jonas Kuhn and Sabine Schulte im Walde	192
<i>Why Find the Right One?</i> Payal Khullar	203