

ACL-IJCNLP 2021

**The 59th Annual Meeting of the Association for  
Computational Linguistics and the 11th International Joint  
Conference on Natural Language Processing**

**Proceedings of the Student Research Workshop**

August 5-6, 2021  
Bangkok, Thailand (online)

©2021 The Association for Computational Linguistics  
and The Asian Federation of Natural Language Processing

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](mailto:acl@aclweb.org)

ISBN 978-1-954085-55-8

# Introduction

Welcome to the ACL-IJCNLP 2021 Student Research Workshop!

The ACL-IJCNLP 2021 Student Research Workshop (SRW) is a forum for student researchers in computational linguistics and natural language processing. The workshop provides a unique opportunity for student participants to present their work and receive valuable feedback from the international research community as well as from faculty mentors.

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, Masters students, and advanced undergraduates 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 feedback on their proposal and ideas about future directions for their work.

This year, the student research workshop has again received wide attention. We received 114 submissions including 109 research papers and 5 thesis proposals. The submissions included 68 long papers and 46 short papers. Following withdrawals and desk rejects, 45 were accepted for an acceptance rate of 39%. Excluding non-archival papers, 36 papers appear in these proceedings. All the accepted papers will be presented virtually in three sessions during the course of August 3rd.

Mentoring is at the heart of the SRW. In keeping with previous years, we had a pre-submission mentoring program before the submission deadline. A total of 36 papers participated in the pre-submission mentoring program. This program offered students the opportunity to receive comments from an experienced researcher to improve the writing style and presentation of their submissions.

We are deeply grateful to the Swiss National Science Foundation (SNSF) for providing funds that covered student registrations. We thank our program committee members for their careful reviews of each paper and all of our mentors for donating their time to provide feedback to our student authors. Thank you to our faculty advisors, Jing Jiang, Rico Sennrich, Derek F. Wong and Nianwen Xue, for their essential advice and guidance, and to the ACL-IJCNLP 2021 organizing committee for their support. Finally, thank you to our student participants!



**Organizers:**

Jad Kabbara, McGill University and the Montreal Institute for Learning Algorithms (MILA)  
Haitao Lin, Institute of Automation, Chinese Academy of Sciences  
Amandalynne Paullada, University of Washington  
Jannis Vamvas, University of Zurich

**Faculty Advisors:**

Jing Jiang, Singapore Management University  
Rico Sennrich, University of Edinburgh  
Derek F. Wong, University of Macau  
Nianwen Xue, Brandeis University

**Pre-submission Mentors:**

Duygu Ataman, University of Zürich  
Valerio Basile, University of Turin  
Eduardo Blanco, University of North Texas  
David Chiang, University of Notre Dame  
Marta R. Costa-Jussà, Universitat Politècnica de Catalunya  
Lucia Donatelli, Saarland University  
Greg Durrett, UT Austin  
Sarah Ebling, University of Zurich  
Yansong Feng, Peking University  
Orhan Firat, Google AI  
Lea Frermann, Melbourne University  
Shujian Huang, National Key Laboratory for Novel Software Technology, Nanjing University  
Kentaro Inui, Tohoku University / Riken  
Robin Jia, Facebook AI Research  
Katharina Kann, University of Colorado Boulder  
Mamoru Komachi, Tokyo Metropolitan University  
Parisa Kordjamshidi, Michigan State University  
Jindřich Libovický, Ludwig Maximilian University of Munich  
Pengfei Liu, Carnegie Mellon University  
Vincent Ng, University of Texas at Dallas  
Sai Krishna Rallabandi, Carnegie Mellon University  
Masoud Rouhizadeh, Johns Hopkins University  
Dipti Sharma, IIIT, Hyderabad  
Manish Shrivastava, International Institute of Information Technology Hyderabad  
Sunayana Sitaram, Microsoft Research India  
Gabriel Stanovsky, The Hebrew University of Jerusalem  
Amanda Stent, Bloomberg  
Hanna Suominen, The Australian National University, Data61/CSIRO, and University of Turku  
Mihai Surdeanu, University of Arizona  
Masashi Toyoda, The University of Tokyo  
Chen-Tse Tsai, Bloomberg LP  
Bonnie Webber, University of Edinburgh  
Yujiu Yang, tsinghua.edu.cn  
Arkaitz Zubiaga, Queen Mary University of London

## Program Committee:

Assina Abdussaitova, Suleyman Demirel University  
Ibrahim Abu Farha, University of Edinburgh  
Oshin Agarwal, University of Pennsylvania  
Piush Aggarwal, University of Duisburg-Essen, Language Technology Lab  
Roei Aharoni, Google  
Miguel A. Alonso, Universidade da Coruña  
Malik Altakrori, McGill University /Mila  
Rami Aly, University of Cambridge  
Bharat Ram Ambati, Apple Inc.  
Aida Amini, University of Washington  
Maria Antoniak, Cornell University  
Tal August, University of Washington  
Vidhisha Balachandran, Carnegie Mellon University  
Anusha Balakrishnan, Microsoft Semantic Machines  
Jorge Balazs, Amazon  
Roberto Basili, University of Roma, Tor Vergata  
Rachel Bawden, Inria  
Chris Biemann, Universität Hamburg  
Tatiana Bladier, Heinrich Heine University Düsseldorf  
Nikolay Bogoychev, University of Edinburgh  
Avishek Joey Bose, Mila/McGill  
Ruken Cakici, METU  
Ronald Cardenas, University of Edinburgh  
Arlene Casey, University of Edinburgh  
Aishik Chakraborty, McGill University  
Jonathan P. Chang, Cornell University  
Jifan Chen, UT Austin  
Sihao Chen, University of Pennsylvania  
Elizabeth Clark, University of Washington  
Xiang Dai, University of Copenhagen  
Siddharth Dalmia, Carnegie Mellon University  
Samvit Dammalapati, Indian Institute of Technology Delhi  
Alok Debnath, Factmata  
Louise Deléger, INRAE - Université Paris-Saclay  
Pieter Delobelle, KU Leuven, Department of Computer Science  
Dorottya Demszky, Stanford University  
Etienne Denis, McGill  
Chris Develder, Ghent University  
Anne Dirkson, Leiden University  
Radina Dobreva, University of Edinburgh  
Zi-Yi Dou, UCLA  
Hicham El Boukkouri, LIMSI, CNRS, Université Paris-Saclay  
Carlos Escolano, Universitat Politècnica de Catalunya  
Luis Espinosa Anke, Cardiff University  
Tina Fang, University of Waterloo  
Murhaf Fares, University of Oslo  
Amir Feder, Technion - Israel Institute of Technology  
Jared Fernandez, Carnegie Mellon University

Dayne Freitag, SRI International  
 Daniel Fried, UC Berkeley  
 Yoshinari Fujinuma, University of Colorado Boulder  
 David Gaddy, University of California, Berkeley  
 Diana Galvan-Sosa, RIKEN AIP  
 Marcos Garcia, Universidade de Santiago de Compostela  
 Arijit Ghosh Chowdhury, Manipal Institute of Technology  
 Liane Guillou, The University of Edinburgh  
 Sarah Gupta, University of Washington  
 Hardy Hardy, The University of Sheffield  
 Mareike Hartmann, University of Copenhagen  
 Junxian He, Carnegie Mellon University  
 Jack Hessel, Allen AI  
 Christopher Homan, Rochester Institute of Technology  
 Junjie Hu, Carnegie Mellon University  
 Jeff Jacobs, Columbia University  
 Aaron Jaech, Facebook  
 Labiba Jahan, Florida International University  
 Tomoyuki Kajiwara, Ehime University  
 Zara Kancheva, IICT-BAS  
 Sudipta Kar, Amazon Alexa AI  
 Alina Karakanta, Fondazione Bruno Kessler (FBK), University of Trento  
 Najoung Kim, Johns Hopkins University  
 Philipp Koehn, Johns Hopkins University  
 Allison Koenecke, Stanford University  
 Mandy Korpusik, Loyola Marymount University  
 Jonathan K. Kummerfeld, University of Michigan  
 Kemal Kurniawan, University of Melbourne  
 Yash Kumar Lal, Stony Brook University  
 Ian Lane, Carnegie Mellon University  
 Alexandra Lavrentovich, Amazon Alexa  
 Lei Li, Peking University  
 Yiyuan Li, University of North Carolina, Chapel Hill  
 Jasy Suet Yan Liew, School of Computer Sciences, Universiti Sains Malaysia  
 Lucy Lin, University of Washington  
 Kevin Lin, Microsoft  
 Fangyu Liu, University of Cambridge  
 Di Lu, Dataminr  
 Chunchuan Lyu, The University of Edinburgh  
 Debanjan Mahata, Bloomberg  
 Valentin Malykh, Huawei Noah's Ark Lab / Kazan Federal University  
 Emma Manning, Georgetown University  
 Courtney Mansfield, University of Washington  
 Pedro Henrique Martins, Instituto de Telecomunicações, Instituto Superior Técnico  
 Bruno Martins, IST and INESC-ID  
 Rui Meng, University of Pittsburgh  
 Antonio Valerio Miceli Barone, The University of Edinburgh  
 Tsvetomila Mihaylova, Instituto de Telecomunicações  
 Farjana Sultana Mim, Tohoku University  
 Sewon Min, University of Washington  
 Koji Mineshima, Keio University

Gosse Minnema, University of Groningen  
 Amita Misra, IBM  
 Omid Moradiannasab, Saarland University  
 Nora Muheim, University of Bern  
 Masaaki Nagata, NTT Corporation  
 Aakanksha Naik, Carnegie Mellon University  
 Denis Newman-Griffis, University of Pittsburgh  
 Dat Quoc Nguyen, VinAI Research  
 Vincent Nguyen, Australian National University & CSIRO Data61  
 Shinji Nishimoto, CiNet  
 Yasumasa Onoe, The University of Texas at Austin  
 Silviu Oprea, University of Edinburgh  
 Naoki Otani, Carnegie Mellon University  
 Ashwin Paranjape, Stanford University  
 Archita Pathak, University at Buffalo (SUNY)  
 Viviana Patti, University of Turin, Dipartimento di Informatica  
 Siyao Peng, Georgetown University  
 Ian Porada, Mila, McGill University  
 Jakob Prange, Georgetown University  
 Adithya Pratapa, Carnegie Mellon University  
 Yusu Qian, New York University  
 Long Qiu, Onehome (Beijing) Network Technology Co. Ltd.  
 Ivaylo Radev, IICT-BAS  
 Sai Krishna Rallabandi, Carnegie Mellon University  
 Vikas Raunak, Microsoft  
 Lina M. Rojas Barahona, Orange Labs  
 Guy Rotman, Faculty of Industrial Engineering and Management, Technion, IIT  
 Maria Ryskina, Carnegie Mellon University  
 Farig Sadeque, Educational Testing Service  
 Jin Sakuma, University of Tokyo  
 Elizabeth Salesky, Johns Hopkins University  
 Younes Samih, University of Düsseldorf  
 Ramon Sanabria, The University Of Edinburgh  
 Michael Sejr Schlichtkrull, University of Amsterdam  
 Sebastian Schuster, New York University  
 Olga Seminck, CNRS  
 Indira Sen, GESIS  
 Vasu Sharma, Carnegie Mellon University  
 Sina Sheikholeslami, KTH Royal Institute of Technology  
 A.B. Siddique, University of California, Riverside  
 Kevin Small, Amazon  
 Marco Antonio Sobrevilla Cabeza, University of São Paulo  
 Katira Soleymanzadeh, Ege University  
 Swapna Somasundaran, Educational Testing Service  
 Sandeep Soni, Georgia Institute of Technology  
 Richard Sproat, Google, Japan  
 Makeesh Narsimhan Sreedhar, Mila, Université de Montréal  
 Tejas Srinivasan, Microsoft  
 Vamshi Krishna Srirangam, International Institute of Information Technology, Hyderabad  
 Marija Stanojevic, Center for Data Analytics and Biomedical Informatics, Temple University  
 Shane Steinert-Threlkeld, University of Washington



Alane Suhr, Cornell University  
Shabnam Tafreshi, The George Washington University  
Wenyi Tay, RMIT University  
Uthayasanker Thayasivam, University of Moratuwa  
Trang Tran, Institute for Creative Technologies, University of Southern California  
Sowmya Vajjala, National Research Council  
Emiel van Miltenburg, Tilburg University  
Dimitrova Vania, University of Leeds  
Rob Voigt, Northwestern University  
Ivan Vulić, University of Cambridge  
Adina Williams, Facebook, Inc.  
Jiacheng Xu, University of Texas at Austin  
Yumo Xu, University of Edinburgh  
Rongtian Ye, Aalto University  
Olga Zamaraeva, University of Washington  
Meishan Zhang, Tianjin University, China  
Justine Zhang, Cornell University  
Ben Zhang, NYU Langone  
Shiyue Zhang, The University of North Carolina at Chapel Hill  
Ben Zhou, University of Pennsylvania  
Zhong Zhou, Carnegie Mellon University



## Table of Contents

<i>Investigation on Data Adaptation Techniques for Neural Named Entity Recognition</i> Evgeniia Tokarchuk, David Thulke, Weiyue Wang, Christian Dugast and Hermann Ney . . . . .	1
<i>Stage-wise Fine-tuning for Graph-to-Text Generation</i> Qingyun Wang, Semih Yavuz, Xi Victoria Lin, Heng Ji and Nazneen Rajani . . . . .	16
<i>Transformer-Based Direct Hidden Markov Model for Machine Translation</i> Weiyue Wang, Zijian Yang, Yingbo Gao and Hermann Ney . . . . .	23
<i>AutoRC: Improving BERT Based Relation Classification Models via Architecture Search</i> Wei Zhu . . . . .	33
<i>How Low is Too Low? A Computational Perspective on Extremely Low-Resource Languages</i> Rachit Bansal, Himanshu Choudhary, Ravneet Punia, Niko Schenk, Émilie Pagé-Perron and Jacob Dahl . . . . .	44
<i>On the Relationship between Zipf’s Law of Abbreviation and Interfering Noise in Emergent Languages</i> Ryo Ueda and Koki Washio . . . . .	60
<i>Long Document Summarization in a Low Resource Setting using Pretrained Language Models</i> Ahsaas Bajaj, Pavitra Dangati, Kalpesh Krishna, Pradhiksha Ashok Kumar, Rheeja Upaal, Bradford Windsor, Eliot Brenner, Dominic Dotterer, Rajarshi Das and Andrew McCallum . . . . .	71
<i>Attending Self-Attention: A Case Study of Visually Grounded Supervision in Vision-and-Language Transformers</i> Jules Samaran, Noa Garcia, Mayu Otani, Chenhui Chu and Yuta Nakashima . . . . .	81
<i>Video-guided Machine Translation with Spatial Hierarchical Attention Network</i> Weiqi Gu, Haiyue Song, Chenhui Chu and Sadao Kurohashi . . . . .	87
<i>Stylistic approaches to predicting Reddit popularity in diglossia</i> Huikai Chua . . . . .	93
<i>"I've Seen Things You People Wouldn't Believe": Hallucinating Entities in GuessWhat?!</i> Alberto Testoni and Raffaella Bernardi . . . . .	101
<i>How do different factors Impact the Inter-language Similarity? A Case Study on Indian languages</i> Sourav Kumar, Salil Aggarwal, Dipti Misra Sharma and Radhika Mamidi . . . . .	112
<i>COVID-19 and Misinformation: A Large-Scale Lexical Analysis on Twitter</i> Dimosthenis Antypas, Jose Camacho-Collados, Alun Preece and David Rogers . . . . .	119
<i>Situation-Based Multiparticipant Chat Summarization: a Concept, an Exploration-Annotation Tool and an Example Collection</i> Anna Smirnova, Evgeniy Slobodkin and George Chernishev . . . . .	127
<i>Modeling Text using the Continuous Space Topic Model with Pre-Trained Word Embeddings</i> Seiichi Inoue, Taichi Aida, Mamoru Komachi and Manabu Asai . . . . .	138
<i>Semantics of the Unwritten: The Effect of End of Paragraph and Sequence Tokens on Text Generation with GPT2</i> He Bai, Peng Shi, Jimmy Lin, Luchen Tan, Kun Xiong, Wen Gao, Jie Liu and Ming Li . . . . .	148

<i>Data Augmentation with Unsupervised Machine Translation Improves the Structural Similarity of Cross-lingual Word Embeddings</i>	
Sosuke Nishikawa, Ryokan Ri and Yoshimasa Tsuruoka . . . . .	163
<i>Joint Detection and Coreference Resolution of Entities and Events with Document-level Context Aggregation</i>	
Samuel Kriman and Heng Ji . . . . .	174
<i>"Hold on honey, men at work": A semi-supervised approach to detecting sexism in sitcoms</i>	
Smriti Singh, Tanvi Anand, Arijit Ghosh Chowdhury and Zeerak Waseem . . . . .	180
<i>Observing the Learning Curve of NMT Systems With Regard to Linguistic Phenomena</i>	
Patrick Stadler, Vivien Macketanz and Eleftherios Avramidis . . . . .	186
<i>Improving the Robustness of QA Models to Challenge Sets with Variational Question-Answer Pair Generation</i>	
Kazutoshi Shinoda, Saku Sugawara and Akiko Aizawa . . . . .	197
<i>Tools Impact on the Quality of Annotations for Chat Untangling</i>	
Jhonny Cerezo, Felipe Bravo-Marquez and Alexandre Henri Bergel . . . . .	215
<i>How Many Layers and Why? An Analysis of the Model Depth in Transformers</i>	
Antoine Simoulin and Benoit Crabbé . . . . .	221
<i>Edit Distance Based Curriculum Learning for Paraphrase Generation</i>	
Sora Kadotani, Tomoyuki Kajiwara, Yuki Arase and Makoto Onizuka . . . . .	229
<i>Changing the Basis of Contextual Representations with Explicit Semantics</i>	
Tamás Ficsor and Gábor Berend . . . . .	235
<i>Personal Bias in Prediction of Emotions Elicited by Textual Opinions</i>	
Piotr Milkowski, Marcin Gruza, Kamil Kanclerz, Przemyslaw Kazienko, Damian Grimling and Jan Kocon . . . . .	248
<i>MVP-BERT: Multi-Vocab Pre-training for Chinese BERT</i>	
Wei Zhu . . . . .	260
<i>CMTA: COVID-19 Misinformation Multilingual Analysis on Twitter</i>	
Raj Pranesh, Mehrdad Farokhenajd, Ambesh Shekhar and Genoveva Vargas-Solar . . . . .	270
<i>Predicting pragmatic discourse features in the language of adults with autism spectrum disorder</i>	
Christine Yang, Duanchen Liu, Qingyun Yang, Zoey Liu and Emily Prud'hommeaux . . . . .	284
<i>SumPubMed: Summarization Dataset of PubMed Scientific Articles</i>	
Vivek Gupta, Perna Bharti, Pegah Nokhiz and Harish Karnick . . . . .	292
<i>A Case Study of Analysis of Construals in Language on Social Media Surrounding a Crisis Event</i>	
Lolo Aboufoul, Khyati Mahajan, Tiffany Gallicano, Sara Levens and Samira Shaikh . . . . .	304
<i>Cross-lingual Evidence Improves Monolingual Fake News Detection</i>	
Daryna Dementieva and Alexander Panchenko . . . . .	310
<i>Neural Machine Translation with Synchronous Latent Phrase Structure</i>	
Shintaro Harada and Taro Watanabe . . . . .	321

<i>Zero Pronouns Identification based on Span prediction</i>	
Sei Iwata, Taro Watanabe and Masaaki Nagata . . . . .	331
<i>On the differences between BERT and MT encoder spaces and how to address them in translation tasks</i>	
Raúl Vázquez, Hande Celikkanat, Mathias Creutz and Jörg Tiedemann . . . . .	337
<i>Synchronous Syntactic Attention for Transformer Neural Machine Translation</i>	
Hiroyuki Deguchi, Akihiro Tamura and Takashi Ninomiya . . . . .	348