

COLING

**International Conference on
Computational Linguistics**

Proceedings of the Conference and Workshops

COLING

Volume 29 (2022), No. 8

**Proceedings of 3rd Workshop on Computational Approaches
to Discourse
(CODI 2022)**

**The 29th International Conference on
Computational Linguistics**

October 16 - 17, 2022
Gyeongju, Republic of Korea

Copyright of each paper stays with the respective authors (or their employers).

ISSN 2951-2093

Preface

Welcome to the 3rd Workshop on Computational Approaches to Discourse, CODI! CODI provides a venue to bring together researchers working on all aspects of discourse in Computational Linguistics and NLP. Our aim is to provide a venue for the entire discourse processing community where we can present and exchange our theories, algorithms, software, datasets, and tools.

The workshop consists of invited talks, contributed papers and extended abstracts. We received paper submissions that span a wide range of topics, addressing issues related to discourse representation and parsing, reference and coreference resolution, dialogue, pragmatics, applications, and more. As the workshop is hybrid this year, papers are presented live either in person or remotely and discussed during live Q&A sessions.

We are happy that CODI 2022 features the CODI-CRAC 2022 shared task on anaphora, bridging, and discourse deixis resolution in dialogue, which is a joint shared task between the CODI and CRAC workshops at COLING. This shared task goes beyond previous ones as its input is transcribed spoken dialogue and it includes anaphoric relations beyond coreference. This effort is spearheaded by Carolyn Rosé. As we hope that the next CODI workshops will also feature shared tasks and other special events, the workshop also includes a discussion on future shared tasks, special sessions on discourse representation and parsing, coreference resolution, and multilingual discourse processing and machine translation.

We thank our invited speakers, **Nate Chambers**, United States Naval Academy, who works on understanding events in text, the semantic relations between them, and how they characterize knowledge they expressed in narratives; and **Bonnie Webber**, University of Edinburgh, known for many decades of seminal research on discourse from early work on “cooperative question-answering” to extended research on discourse anaphora and discourse relations. We would also like to thank our reviewers for their thoughtful and instructive comments. They helped us to prepare an inclusive workshop program. Finally we would like to thank the COLING 2022 workshop chairs Sadao Kurohashi, Patrizia Paggio and Nianwen Xue who organized the COLING workshop program.

The CODI Organizers,

Chloé Braud, Christian Hardmeier, Junyi Jessy Li, Sharid Loáiciga, Michael Strube, and Amir Zeldes

Organizers:

Chloé Braud, CNRS - IRIT
Christian Hardmeier, IT University of Copenhagen and Uppsala University¹
Junyi Jessy Li, The University of Texas at Austin
Sharid Loáiciga, University of Gothenburg
Michael Strube, Heidelberg Institute for Theoretical Studies
Amir Zeldes, Georgetown University

Program Committee:

Malihe Alikhani, University of Pittsburgh, US
Giuseppe Carenini, University of British Columbia, Canada
Haixia Chai, Heidelberg Institute for Theoretical Studies, Germany
Jackie Chi Kit Cheung, McGill University, Canada
Iria da Cunha, Universidad Nacional de Educación a Distancia, Spain
Vera Demberg, Saarland University, Germany
Pascal Denis, INRIA, France
Stefanie Dipper, Ruhr-Universität Bochum, Germany
Elisa Ferracane, Abridge, US
Zhengxian Gong, Soochow University, China
Yulia Grishina, Amazon, Germany
Ryuichiro Higashinaka, NTT, Japan
Ruihong Huang, Texas A&M University, US
Kentaro Inui, Tohoku University, Japan
Yangfeng Ji, University of Virginia, US
Yaman Kumar, IIIT Delhi, India
Murathan Kurfalı, Stockholm University, Sweden
Sadao Kurohashi, Kyoto University, Japan
Sobha Lalitha Devi, AU-KBC Research Centre, India
Ekaterina Lapshinova-Koltunski, Saarland University, Germany
Philippe Muller, IRIT, France
Mark-Christoph Müller, Heidelberg Institute for Theoretical Studies, Germany
Anna Nedoluzhko, Charles University, Czechia
Vincent Ng, University of Texas at Dallas, US
Maciej Ogrodniczuk, Polish Academy of Sciences, Poland
Thiago Pardo, University of São Paulo, Brazil
Massimo Poesio, Queen Mary University of London, UK
Hannah Rohde, University of Edinburgh, UK
Attapol Rutherford, Chulalongkorn University, Thailand
Manfred Stede, University of Potsdam, Germany
Francielle Vargas, University of São Paulo, Brazil
Bonnie Webber, University of Edinburgh, UK
Deyi Xiong, Tianjin University, China
Hongfei Xu, Saarland University, Germany
Nianwen Xue, Brandeis University, US
Heike Zinsmeister, University of Hamburg, Germany

Invited Speakers:

Nate Chambers, United States Naval Academy, US
Bonnie Webber, University of Edinburgh, UK

¹Christian Hardmeier was supported by the Swedish Research Council under grant 2017-930.

Table of Contents

<i>KOJAK: A New Corpus for Studying German Discourse Particle ja</i> Adil Soubki, Owen Rambow and Chong Kang	1
<i>Improving Topic Segmentation by Injecting Discourse Dependencies</i> Linzi Xing, Patrick Huber and Giuseppe Carenini	7
<i>Evaluating How Users Game and Display Conversation with Human-Like Agents</i> Won Ik Cho, Soomin Kim, Eujeong Choi and Younghoon Jeong	19
<i>Evaluating Discourse Cohesion in Pre-trained Language Models</i> Jie He, Wanqiu Long and Deyi Xiong	28
<i>Easy-First Bottom-Up Discourse Parsing via Sequence Labelling</i> Andrew Shen, Fajri Koto, Jey Han Lau and Timothy Baldwin	35
<i>Using Translation Process Data to Explore Explicitation and Implication through Discourse Connectives</i> Ekaterina Lapshinova-Koltunski and Michael Carl	42
<i>Label distributions help implicit discourse relation classification</i> Frances Yung, Kaveri Anuranjana, Merel Scholman and Vera Demberg	48
<i>The Keystone Role Played by Questions in Debate</i> Zlata Kikteva, Kamila Gorska, Wassiliki Siskou, Annette Hautli-Janisz and Chris Reed	54
<i>Shallow Discourse Parsing for Open Information Extraction and Text Simplification</i> Christina Niklaus, André Freitas and Siegfried Handschuh	64
<i>Predicting Political Orientation in News with Latent Discourse Structure to Improve Bias Understanding</i> Nicolas Devatine, Philippe Muller and Chloé Braud	77
<i>Attention Modulation for Zero-Shot Cross-Domain Dialogue State Tracking</i> Mathilde Veron, Olivier Galibert, Guillaume Bernard and Sophie Rosset	86
<i>An Empirical Study of Topic Transition in Dialogue</i> Mayank Soni, Brendan Spillane, Leo Muckley, Orla Cooney, Emer Gilmartin, Christian Saam, Benjamin Cowan and Vincent Wade	92

Conference Program

October 16, 2022

9:15–9:30 *Opening Remarks*

9:30–10:30 *Invited Talk*
Nate Chambers

10:30–11:00 *Coffee Break*

11:00–11:10 *KOJAK: A New Corpus for Studying German Discourse Particle ja*
Adil Soubki, Owen Rambow and Chong Kang

11:10–11:20 *Discourse Analysis via Questions and Answers: Parsing Dependency Structures of Questions Under Discussion*
Wei-Jen Ko, Dananjay Srinivas, Cutter Dalton, Greg Durrett and Junyi Jessy Li

11:20–11:30 *Towards Domain-Independent Supervised Discourse Parsing Through Gradient Boosting*
Patrick Huber and Giuseppe Carenini

11:30–11:45 *Improving Topic Segmentation by Injecting Discourse Dependencies*
Linzi Xing, Patrick Huber and Giuseppe Carenini

11:45–11:55 *Evaluating How Users Game and Display Conversation with Human-Like Agents*
Won Ik Cho, Soomin Kim, Eujeong Choi and Younghoon Jeong

11:55–12:05 *Evaluating Discourse Cohesion in Pre-trained Language Models*
Jie He, Wanqiu Long and Deyi Xiong

12:05–12:15 *Easy-First Bottom-Up Discourse Parsing via Sequence Labelling*
Andrew Shen, Fajri Koto, Jey Han Lau and Timothy Baldwin

12:15–12:30 *Open discussion on future CODI workshops and shared tasks*

12:30–14:00 *Lunch*

October 16, 2022 (continued)

- 14:00–14:10 *Multi-Party Conversation Generation via Discourse Structure Planning*
Alexander Chernyavskiy and Dmitry Ilvovsky
- 14:10–14:20 *Explicitation Phenomena and Surprisal Effects in Translation and Interpreting*
Ekaterina Lapshinova-Koltunski, Christina Pollkläsener and Heike Przybyl
- 14:20–14:30 *Using Translation Process Data to Explore Explicitation and Implication through Discourse Connectives*
Ekaterina Lapshinova-Koltunski and Michael Carl
- 14:30–14:45 *Towards Identifying Alternative-Lexicalization Signals of Discourse Relations*
René Knaebel and Manfred Stede
- 14:45–14:55 *Towards Semi-Supervised Discourse Graph Parsing in Dialogs with Pre-trained Language Models and Linguistic Features*
Chuyuan Li, Patrick Huber, Wen Xiao, Maxime Amblard, Chloé Braud and Giuseppe Carenini
- 14:55–15:05 *Label distributions help implicit discourse relation classification*
Frances Yung, Kaveri Anuranjana, Merel Scholman and Vera Demberg
- 15:05–15:15 *Bullshit and Unanswerable QUDs*
Oliver Deck
- 15:15–15:30 *The Keystone Role Played by Questions in Debate*
Zlata Kikteva, Kamila Gorska, Wassiliki Siskou, Annette Hautli-Janisz and Chris Reed
- 15:30–16:00 *Coffee Break***
- 16:00–16:15 *Shallow Discourse Parsing for Open Information Extraction and Text Simplification*
Christina Niklaus, André Freitas and Siegfried Handschuh
- 16:15–16:25 *Predicting Political Orientation in News with Latent Discourse Structure to Improve Bias Understanding*
Nicolas Devatine, Philippe Muller and Chloé Braud
- 16:25–16:35 *Attention Modulation for Zero-Shot Cross-Domain Dialogue State Tracking*
Mathilde Veron, Olivier Galibert, Guillaume Bernard and Sophie Rosset
- 16:35–16:45 *An Empirical Study of Topic Transition in Dialogue*
Mayank Soni, Brendan Spillane, Leo Muckley, Orla Cooney, Emer Gilmartin, Christian Saam, Benjamin Cowan and Vincent Wade

October 16, 2022 (continued)

16:45–17:45 *Invited Talk*
Bonnie Webber

17:45–18:00 *Open discussion on future CODI workshops and shared tasks*

October 17, 2022

CODI-CRAC Shared Task

14:00 **Session 1: Shared Task Competition**

14:00 *Welcome*

14:05 *The CODI-CRAC 2022 Shared Task on Anaphora, Bridging, and Discourse Deixis in Dialogue*
Juntao Yu, Sopan Khosla, Ramesh Manuvinakurike, Lori Levin, Vincent Ng, Massimo Poesio, Michael Strube and Carolyn Rosé

14:35 *Anaphora Resolution in Dialogue: System Description (CODI-CRAC 2022 Shared Task)*
Tatiana Anikina, Natalia Skachkova, Joseph Renner and Priyansh Trivedi

14:50 *Pipeline Coreference Resolution for Anaphoric Identity in Dialogues*
Damrin Kim, Seongsik Park, Mirae Han and Harksoo Kim

15:05 *Neural Anaphora Resolution in Dialogue Revisited*
Shengjie Li, Hideo Kobayashi and Vincent Ng

15:30 *Coffee Break*

October 17, 2022 (continued)

16:00 **Session 2: Keynote Session**

16:00 *Invited Talk: Massimo Poesio and Lori Levin*

16:45 *Open Discussion*

17:45 *Closing Remarks*