

NAACL HLT 2019

**The Workshop on Discourse Relation  
Parsing and Treebanking**

**Proceedings of the Workshop**

June 6, 2019  
Minneapolis, MN

©2019 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](mailto:acl@aclweb.org)

ISBN 978-1-948087-98-8

**Proceedings of DISRPT 2019 -  
The Workshop on Discourse Relation Parsing and Treebanking**

Edited by:

Amir Zeldes, Georgetown University

Debopam Das, Humboldt University of Berlin

Erick Galani Maziero, University of São Paulo

Juliano Desiderato Antonio, Universidade Estadual de Maringa

Mikel Iruskieta, University of the Basque Country



**Organizers:**

Amir Zeldes, Georgetown University  
Debopam Das, Humboldt University of Berlin  
Erick Galani Maziero, University of São Paulo  
Juliano Desiderato Antonio, Universidade Estadual de Maringa  
Mikel Iruskietta, University of the Basque Country

**Program Committee:**

Stergos Afantenos, IRIT - Université Paul Sabatier, France  
Farah Benamara, IRIT - Université Paul Sabatier, France  
Eduard Hovy, Carnegie Mellon University, USA  
Irene Castellon, Universitat de Barcelona, Spain  
Johann Christian Chiarcos, Wolfgang Goethe Universität Frankfurt, Germany  
Maria Beatriz Nascimento Decat, Universidade Federal de Minas Gerais, Brazil  
Iria da Cunha, Universidad Nacional de Educación a Distancia, Spain  
Barbara Di Eugenio, University of Illinois at Chicago, USA  
Arantza Diaz de Ilarraza, University of the Basque Country, Spain  
Flavius Frasincaer, Erasmus University Rotterdam, Netherlands  
Maria Eduarda Giering, Universidade do Vale do Rio dos Sinos, Brazil  
Nancy Green, University of North Carolina, USA  
Kerstin Kunz, Universität Heidelberg, Germany  
Ekaterina Lapshinova-Koltunski, Universität des Saarlandes, Germany  
Jiri Mirovsky, Charles University, Czech Republic  
Anna Nedoluzhko, Charles University, Czech Republic  
Thiago Pardo, Universidade de São Paulo, Brazil  
Lucie Polakova, Charles University, Czech Republic  
Gisela Redeker, University of Groningen, Netherlands  
Hannah Rohde, University of Edinburgh, UK  
Gerardo Sierra, Universidad Nacional Autónoma de México, Mexico  
Christian Stab, Technische Universität Darmstadt, Germany  
Manfred Stede, Universität Potsdam, Germany  
Mihai Surdeanu, University of Arizona, USA  
Maite Taboada, Simon Fraser University, Canada  
Juan-Manuel Torres, Laboratoire Informatique d'Avignon, France  
Nianwen Xue, Brandeis University, USA

**Invited Speaker:**

Bonnie Webber, University of Edinburgh



## Table of Contents

<i>Introduction to Discourse Relation Parsing and Treebanking (DISRPT): 7th Workshop on Rhetorical Structure Theory and Related Formalisms</i>	
Amir Zeldes, Debopam Das, Erick Galani Maziero, Juliano Antonio and Mikel Iruskieta . . . . .	1
<i>Toward Cross-theory Discourse Relation Annotation</i>	
Peter Bourgonje and Olha Zolotareno . . . . .	7
<i>Acquiring Annotated Data with Cross-lingual Explicitation for Implicit Discourse Relation Classification</i>	
Wei Shi, Frances Yung and Vera Demberg . . . . .	12
<i>From News to Medical: Cross-domain Discourse Segmentation</i>	
Elisa Ferracane, Titan Page, Junyi Jessy Li and Katrin Erk . . . . .	22
<i>Nuclearity in RST and signals of coherence relations</i>	
Debopam Das . . . . .	30
<i>The Rhetorical Structure of Attribution</i>	
Andrew Potter . . . . .	38
<i>Annotating Shallow Discourse Relations in Twitter Conversations</i>	
Tatjana Scheffler, Berfin Aktaş, Debopam Das and Manfred Stede . . . . .	50
<i>A Discourse Signal Annotation System for RST Trees</i>	
Luke Gessler, Yang Liu and Amir Zeldes . . . . .	56
<i>EusDisParser: improving an under-resourced discourse parser with cross-lingual data</i>	
Mikel Iruskieta and Chloé Braud . . . . .	62
<i>Beyond The Wall Street Journal: Anchoring and Comparing Discourse Signals across Genres</i>	
Yang Liu . . . . .	72
<i>Towards the Data-driven System for Rhetorical Parsing of Russian Texts</i>	
Artem Shelmanov, Dina Pisarevskaya, Elena Chistova, Svetlana Toldova, Maria Kobozeva and Ivan Smirnov . . . . .	82
<i>RST-Tace A tool for automatic comparison and evaluation of RST trees</i>	
Shujun Wan, Tino Kutschbach, Anke Lüdeling and Manfred Stede . . . . .	88
<i>The DISRPT 2019 Shared Task on Elementary Discourse Unit Segmentation and Connective Detection</i>	
Amir Zeldes, Debopam Das, Erick Galani Maziero, Juliano Antonio and Mikel Iruskieta . . . . .	97
<i>Multi-lingual and Cross-genre Discourse Unit Segmentation</i>	
Peter Bourgonje and Robin Schäfer . . . . .	105
<i>ToNy: Contextual embeddings for accurate multilingual discourse segmentation of full documents</i>	
Philippe Muller, Chloé Braud and Mathieu Morey . . . . .	115
<i>Multilingual segmentation based on neural networks and pre-trained word embeddings</i>	
Mikel Iruskieta, Kepa Bengoetxea, Aitziber Atutxa Salazar and Arantza Diaz de Ilarraza . . . . .	125
<i>GumDrop at the DISRPT2019 Shared Task: A Model Stacking Approach to Discourse Unit Segmentation and Connective Detection</i>	
Yue Yu, Yilun Zhu, Yang Liu, Yan Liu, Siyao Peng, Mackenzie Gong and Amir Zeldes . . . . .	133

<i>Towards discourse annotation and sentiment analysis of the Basque Opinion Corpus</i> Jon Alkorta, Koldo Gojenola and Mikel Iruskieta .....	144
<i>Using Rhetorical Structure Theory to Assess Discourse Coherence for Non-native Spontaneous Speech</i> Xinhao Wang, Binod Gyawali, James V. Bruno, Hillary R. Molloy, Keelan Evanini and Klaus Zechner .....	153
<i>Applying Rhetorical Structure Theory to Student Essays for Providing Automated Writing Feedback</i> Shiyan Jiang, Kexin Yang, Chandrakumari Suvarna, Pooja Casula, Mingtong Zhang and Carolyn Rose .....	163



# Conference Program

**Thursday, June 6, 2019**

**09:00–12:00** *Session S1: Session 1*

09:00–09:15 *Opening remarks*

Amir Zeldes, Debopam Das, Erick Galani Maziero, Juliano Antonio Desiderato,  
Mikel Iruskieta

09:15–10:15 *Invited Talk: "Discourse (2009–2019): Recent successes, future challenges"*

Bonnie Webber

**10:15–10:30** *Coffee break*

10:30–11:00 *Toward Cross-theory Discourse Relation Annotation*

Peter Bourgonje and Olha Zolotarenko

11:00–11:30 *Acquiring Annotated Data with Cross-lingual Explicitation for Implicit Discourse Relation Classification*

Wei Shi, Frances Yung and Vera Demberg

11:30–12:00 *From News to Medical: Cross-domain Discourse Segmentation*

Elisa Ferracane, Titan Page, Junyi Jessy Li and Katrin Erk

**12:00–13:30** *Lunch break*

**13:30–15:00** *Session S2: Session 2*

13:30–14:00 *Nuclearity in RST and signals of coherence relations*

Debopam Das

14:00–14:30 *The Rhetorical Structure of Attribution*

Andrew Potter

14:30–15:00 *Annotating Shallow Discourse Relations in Twitter Conversations*

Tatjana Scheffler, Berfin Aktaş, Debopam Das and Manfred Stede

**Thursday, June 6, 2019 (continued)**

**15:00–16:00** *Session P1: Regular posters*

*A Discourse Signal Annotation System for RST Trees*

Luke Gessler, Yang Liu and Amir Zeldes

*EusDisParser: improving an under-resourced discourse parser with cross-lingual data*

Mikel Iruskieta and Chloé Braud

*Beyond The Wall Street Journal: Anchoring and Comparing Discourse Signals across Genres*

Yang Liu

*Towards the Data-driven System for Rhetorical Parsing of Russian Texts*

Artem Shelmanov, Dina Pisarevskaya, Elena Chistova, Svetlana Toldova, Maria Kobozeva and Ivan Smirnov

*RST-Tace A tool for automatic comparison and evaluation of RST trees*

Shujun Wan, Tino Kutschbach, Anke Lüdeling and Manfred Stede

**15:00–16:00** *Session P2: Shared task posters*

*Multi-lingual and Cross-genre Discourse Unit Segmentation*

Peter Bourgonje and Robin Schäfer

*ToNy: Contextual embeddings for accurate multilingual discourse segmentation of full documents*

Philippe Muller, Chloé Braud and Mathieu Morey

*Multilingual segmentation based on neural networks and pre-trained word embeddings*

Mikel Iruskieta, Kepa Bengoetxea, Aitziber Atutxa Salazar and Arantza Diaz de Ilarraza

*GumDrop at the DISRPT2019 Shared Task: A Model Stacking Approach to Discourse Unit Segmentation and Connective Detection*

Yue Yu, Yilun Zhu, Yang Liu, Yan Liu, Siyao Peng, Mackenzie Gong and Amir Zeldes

**16:00–18:15** *Session S3: Session 3*

**Thursday, June 6, 2019 (continued)**

16:00–16:30 *Towards discourse annotation and sentiment analysis of the Basque Opinion Corpus*  
Jon Alkorta, Koldo Gojenola and Mikel Iruskieta

16:30–17:00 *Using Rhetorical Structure Theory to Assess Discourse Coherence for Non-native Spontaneous Speech*  
Xinhao Wang, Binod Gyawali, James V. Bruno, Hillary R. Molloy, Keelan Evanini and Klaus Zechner

17:00–17:30 *Applying Rhetorical Structure Theory to Student Essays for Providing Automated Writing Feedback*  
Shiyan Jiang, Kexin Yang, Chandrakumari Suvarna, Pooja Casula, Mingtong Zhang and Carolyn Rose

**17:30–18:00** *Open meeting*

18:00–18:15 *Closing remarks*  
Amir Zeldes, Debopam Das, Erick Galani Maziero, Juliano Antonio Desiderato, Mikel Iruskieta

