

NAACL HLT 2019

**Second Workshop on Computational Models of
Reference, Anaphora and Coreference**

Proceedings of the Workshop

June 7, 2019
Minneapolis, USA

©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

ISBN 978-1-948087-97-1

Introduction

This is the second edition of the Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC), which was first held in New Orleans last year in conjunction with NAACL HLT 2018. CRAC and its predecessor, the Coreference Resolution Beyond OntoNotes (CORBON) workshop series that started in 2016, have arguably become the primary forum for coreference researchers to present their latest results since the demise of the Discourse Anaphora and Anaphor Resolution Colloquium series in 2011. While CORBON focuses on under-investigated coreference phenomena, CRAC has a broader scope, covering all cases of computational modeling of reference, anaphora, and coreference.

The workshop received 10 submissions: half of them were from Europe, two were from the U.S., and the remaining three were from India. We are pleased to see that the submissions covered not only a variety of less-studied languages in the coreference community (e.g., Basque, French, German, Malayalam or Tamil) but also many under-investigated topics in coreference resolution (e.g., feature representation, coreference for low-resource languages, coreference in specialized domains, and evaluation of coreference resolvers). While it is perhaps not surprising to receive submissions focusing on the design and use of neural models for coreference resolution given the recent popularity of deep learning for natural language processing, it is interesting to see that the most popular topic among the submitted papers is cross-lingual coreference resolution. In fact, one of the workshop sessions will be devoted entirely to this topic.

As in previous years, each submission was rigorously reviewed by three to five programme committee members. We would like to thank the 18 programme committee members for their hard work. Based on their recommendations, we initially accepted four papers and conditionally accepted two papers. Both conditionally accepted papers were eventually accepted to the workshop after we made sure that the authors adequately addressed the reviewers' comments in the final camera-ready version. All of the accepted papers will be presented orally.

We are grateful to Amir Zeldes for accepting our invitation to be this year's invited speaker. Amir will give a talk on coreference, discourse structure and coherence.

Finally, we would like to thank the workshop participants for joining in. We look forward to an exciting workshop in Minneapolis.

— Maciej Ogrodniczuk, Sameer Pradhan, Yulia Grishina, and Vincent Ng

Organizing Committee and Proceedings Editors:

Maciej Ogrodniczuk, Institute of Computer Science, Polish Academy of Sciences
Sameer Pradhan, cemantix.org and Vassar College
Yulia Grishina, Amazon
Vincent Ng, University of Texas at Dallas

Programme Committee:

Antonio Branco, University of Lisbon
Stephanie Dipper, University of Bochum
Yulia Grishina, Amazon
Veronique Hoste, Ghent University
Ryu Iida, National Institute of Information and Communications Technology (NICT)
Sandra Kübler, Indiana University
Sobha Lalitha Devi, AU-KBC Research Center, Anna University of Chennai
Emmanuel Lassalle, Machina Capital, Paris
Katja Markert, Heidelberg University
Pavankumar Reddy Muddireddy, Google
Costanza Navaretta, University of Copenhagen
Anna Nedoluzhko, Charles University in Prague
Michal Novak, Charles University in Prague
Constantin Orasan, University of Wolverhampton
Massimo Poesio, Queen Mary University of London
Marta Recasens, Google
Yannick Versley, IBM
Heike Zinsmeister, University of Hamburg

Invited Speaker:

Amir Zeldes, Georgetown University

Table of Contents

<i>Evaluation of Named Entity Coreference</i>	
Oshin Agarwal, Sanjay Subramanian, Ani Nenkova and Dan Roth	1
<i>Neural Coreference Resolution with Limited Lexical Context and Explicit Mention Detection for Oral French</i>	
Loïc Grobol	8
<i>Entity Decisions in Neural Language Modelling: Approaches and Problems</i>	
Jenny Kunz and Christian Hardmeier	15
<i>Cross-lingual NIL Entity Clustering for Low-resource Languages</i>	
Kevin Blissett and Heng Ji	20
<i>Cross-lingual Incongruences in the Annotation of Coreference</i>	
Ekaterina Lapshinova-Koltunski, Sharid Loáiciga, Christian Hardmeier and Pauline Krielke	26
<i>Deep Cross-Lingual Coreference Resolution for Less-Resourced Languages: The Case of Basque</i>	
Gorka Urbizu, Ander Soraluze and Olatz Arregi	35

Workshop Program: June 7, 2019

09:00–10:30 Session 1: Welcome and Invited Talk

09:00–09:15 *Introduction*

Vincent Ng and Sameer Pradhan

09:15–10:30 Invited talk: *Coreference and Coherence Revisited*

Amir Zeldes

10:30–11:00 Mid-Morning Break

11:00–12:30 Session 2

11:00–11:30 *Evaluation of Named Entity Coreference*

Oshin Agarwal, Sanjay Subramanian, Ani Nenkova and Dan Roth

11:30–12:00 *Neural Coreference Resolution with Limited Lexical Context
and Explicit Mention Detection for Oral French*

Loïc Grobol

12:00–12:30 *Entity Decisions in Neural Language Modelling: Approaches and Problems*

Jenny Kunz and Christian Hardmeier

12:30–14:00 Lunch Break

14:00–15:30 Session 3

14:00–14:30 *Cross-lingual NIL Entity Clustering for Low-resource Languages*

Kevin Blissett and Heng Ji

14:30–15:00 *Cross-lingual Incongruences in the Annotation of Coreference*

Ekaterina Lapshinova-Koltunski, Sharid Loáiciga, Christian Hardmeier
and Pauline Krielke

15:00–15:30 *Deep Cross-Lingual Coreference Resolution for Less-Resourced Languages:
The Case of Basque*

Gorka Urbizu, Ander Soraluze and Olatz Arregi

15:30–16:00 Mid-Afternoon Break

16:00–17:00 Session 4

16:00–17:00 *Informal Discussion and Closing Remarks*

Sameer Pradhan

