

CORBON 2016:
Coreference Resolution Beyond OntoNotes

NAACL HLT 2016 Workshop

Maciej Ogrodniczuk and Vincent Ng

June 16, 2016

Main Goals

- Coreference resolution for **less-researched languages**
- **Non-identity coreference**
 - e.g., bridging references
- **Less-researched phenomena**
 - e.g., ellipsis, split antecedents, reference to abstract entities
- **Difficult cases of anaphora and coreference**
 - e.g. those whose resolution requiring **discourse** and **pragmatic** knowledge

Submissions

- Deadline: March 6
- 20 submissions

Topics

- Question answering
- Event anaphora
- Cataphora, generic anaphors, metonymy
- Bridging
- Abstract entities
- Anaphoric connectives
- VP ellipsis
- Split antecedents
- Annotation projection
- Coreference for NLP applications: Slot filling
- Identify coreference

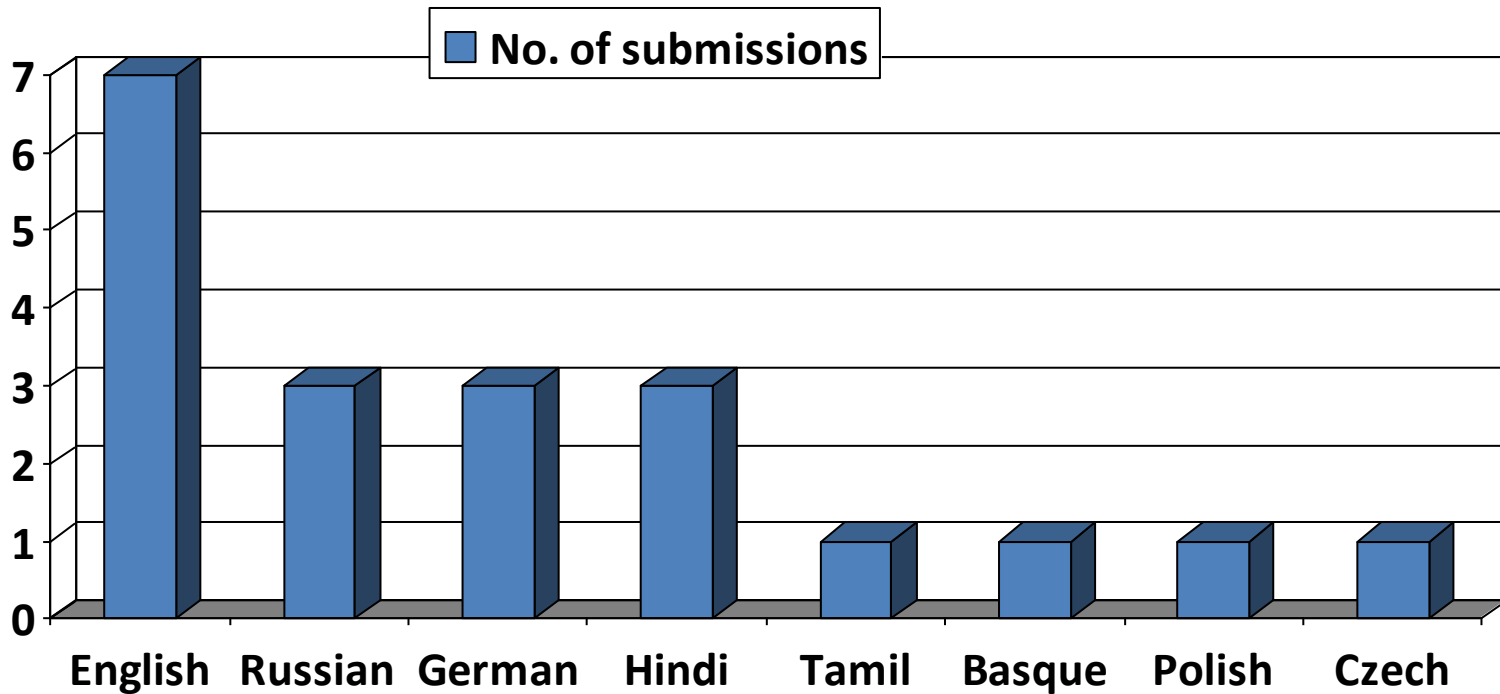
Topics

- Question answering
- Event anaphora
- Cataphora, generic anaphors, metonymy
- Bridging
- Abstract entities
- Anaphoric connectives
- VP ellipsis
- Split antecedents
- Annotation projection
- Coreference for NLP applications: Slot filling
- Identify coreference

Topics

- Question answering
- Event anaphora
- Cataphora, generic anaphors, metonymy
- Bridging
- Abstract entities
- Anaphoric connectives
- VP ellipsis
- Split antecedents
- Annotation projection
- Coreference for NLP applications: Slot filling
- Identify coreference

Submissions: Languages Investigated



Paper Reviewing and Selection

- Each submission was reviewed by three PC members
- Reviewers rated each paper on a seven-point scale
 - 3: Strong accept
 - 2: Accept
 - 1: Weak accept
 - 0: Borderline
 - -1: Weak reject
 - -2: Reject
 - -3: Strong reject

Paper Reviewing and Selection

- Accepted all submissions with a positive average score
- Rejected all submissions with an average score of 0 or below with one exception
 - average score of 0 (+3, 0, -3)
- Accepted 13 of the 20 submissions

Program Committee

Anders Björkelund

Antonio Branco

Dan Cristea

Sobha Lalitha Devi

Lars Hellan

Veronique Hoste

Yufang Hou

Sandra Kübler

Sebastian Martschat

Ruslan Mitkov

Costanza Navaretta

Anna Nedoluzhko

Vincent Ng

Michal Novak

Maciej Ogrodniczuk

Constantin Orasan

Simone Paolo Ponzetto

Massimo Poesio

Sameer Pradhan

Marta Recasens

Agata Savary

Olga Uryupina

Yannick Versley

Desislava Zhekova

Heike Zinsmeister

Shared Task

- **Goal:** To investigate how well one can build a coreference resolver for a "surprise" language for which only a small amount of coreference-annotated data is available for training
- **Motivation:** To promote development of coreference technologies that are applicable to a larger number of languages than currently possible
- **Data:** 350K-token manually annotated corpus for Polish, links to other available Polish data, scorer, documentation
- **Tracks:**
 - Coreference resolution on extracted mentions (CREME)
 - Coreference resolution on plain text (CROP)

Workshop Schedule

- **Session 1 (9:00-10:30)**
 - Invited talk by Michael Strube + 1 short talk
- Morning coffee break (10:30-11:00)
- **Session 2 (11:00-12:30)**
 - 3 long talks
- Lunch break (12:30-14:00)
- **Session 3 (14:00-15:30)**
 - Invited talk by Andy Kehler + 1 long talk
- Afternoon coffee break (15:30-16:00)
- **Session 4 (16:00-17:30)**
 - One-minute madness + poster session

Feedback

- Let us know your feedback during and after the workshop
 - Planning for future workshops
 - JNLE special issue proposal

Announcements

- Oral presenters
 - Email Maciej your slides by the end of the day today
- Poster and 1-minute madness presenters
 - Load your slide(s) into Maciej's laptop during the afternoon coffee break
 - Set up your poster during afternoon coffee break
 - Email Maciej the PDF version of your poster by the end of the day today