

**CORBON 2017:  
Second Workshop on  
Coreference Resolution Beyond OntoNotes**

**EACL 2017 Workshop**

**Maciej Ogrodniczuk and Vincent Ng**

April 4, 2017

# CORBON 2017

- Second workshop in the series
  - First workshop held in conjunction with NAACL 2016 in San Diego
    - Majority of the submissions were from Europe
      - Motivated us to organise CORBON 2017 in Europe

# 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

# Shared Task

- **Goal:** To investigate how well one can build a coreference resolver for a "surprise" language for which no 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
- **Shared task organiser:** Yulia Grishina
- More on the shared task in Session 4

# Panel Discussion

- On the interplay of coreference and discourse
- Cancelled due to panelists unavailability
  - Replaced with invited talks by Ruslan Mitkov and Anna Nedoluzhko

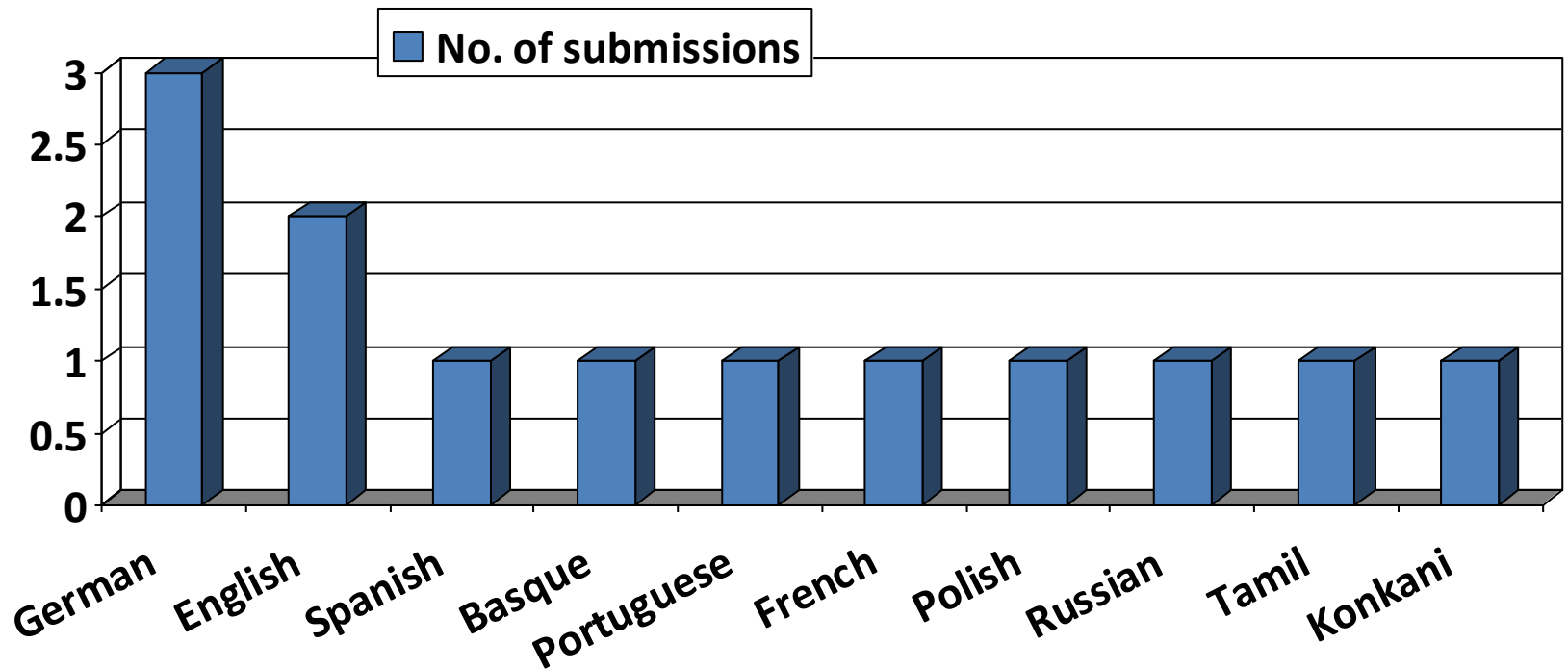
# Submissions

- Deadline: January 23
- 12 submissions
  - Down from 20 in CORBON 2016

# Submissions: Topics

- Coreference for NLP applications: Machine Translation
- Winograd Schema
- Semantic knowledge sources for coreference
- Feature representations
- Annotation projection
- Measuring language and register contrasts
- Resolving pronouns suffixed with relative participle verbs
- Mention detection
- Identify coreference
- Code switching

# Submissions: Languages Investigated





# Paper Reviewing and Selection

- Each submission was reviewed by at least three PC members
  - In most cases, we had 4-5 reviews per submission
- 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
- Accepted 6 of the 12 submissions
- Compared to last year, this year's average score went down
  - Partly due to the increase in number of reviews per submission

# Programme Committee

Anders Björkelund

Antonio Branco

Chen Chen

Dan Cristea

Pascal Denis

Sobha Lalitha Devi

Yulia Grishina

Lars Hellan

Veronique Hoste

Yufang Hou

Ryu Iida

Ekaterina Lapshinova-Koltunski

Emmanuel Lassalle

Sebastian Martschat

Ruslan Mitkov

Costanza Navaretta

Anna Nedoluzhko

Michal Novak

Maciej Ogrodniczuk

Constantin Orasan

Massimo Poesio

Manfred Stede

Veselin Stoyanov

Yannick Versley

Rob Voigt

Sam Wiseman

Amir Zeldes

Heike Zinsmeister

# Workshop Schedule

- **Session 1 (9:30-11:00)**
  - Invited talk by Massimo Poesio
- Morning coffee break (11:00-11:30)
- **Session 2 (11:30-13:00)**
  - 3 oral presentations
- Lunch break (13:00-14:30)
- **Session 3 (14:30-16:00)**
  - Invited talks by Ruslan Mikov and Anna Nedoluzhko
  - 1 oral presentation
- Afternoon coffee break (16:00-16:30)
- **Session 4 (16:30-18:00)**
  - 1 oral presentation
  - Shared task (overview by Yulia + talk on participating system)

# Feedback

- Let us know your feedback during and after the workshop
  - Planning for future workshops

# Announcements

- Schedule change
  - *Enriching Basque Coreference Resolution System using Semantic Knowledge Sources* **cancelled**
  - *Use Generalized Representations, But Do Not Forget Surface Features* by Moosavi & Strube **moved from session 1 to session 2**
- Presenters
  - Email Maciej your slides by the end of the day today