

## Tuesday 23 May: Main Conference

### Room: Wallenbergsalen

08.30– Registration

**09:00–09:30** Opening

**09:30–10:30** Keynote: Kyunghyun Cho **Chair: Jörg Tiedemann**

10:30–11:00 Coffee break

**Session: Syntax** **Chair: Gerlof Bouma**

**11:00–12:30** Joint UD Parsing of Norwegian Bokmål and Nynorsk **Erik Velldal, Lilja Øvrelid and Petter Hohle**

Replacing OOV Words For Dependency Parsing With Distributional Semantics **Prasanth Kolachina, Martin Riedl and Chris Biemann**

Real-valued Syntactic Word Vectors (RSV) for Greedy Neural Dependency Parsing **Ali Basirat and Joakim Nivre**

12:30–13:30 Lunch

**Session: Social media** **Chair: Filip Ginter**

**13:30–14:30** Twitter Topic Modeling by Tweet Aggregation **Asbjørn Steinskog, Jonas Therkelsen and Björn Gambäck**

A Multilingual Entity Linker Using PageRank and Semantic Graphs **Anton Södergren and Pierre Nugues**

**14:30–15:00** Lightning talks **Chair: Jörg Tiedemann**

15:00–15:30 Coffee break

15:30–16:30 Poster session

**16:30–17:15** Tutorial on Swedish: Maia Andréasson

17:15–18:00 SIG-ICALL meeting

19:00– Conference dinner

## Tuesday 23 May: Main Conference

### Room: Europe

08.30– Registration

09:00–09:30 Opening

09:30–10:30 Keynote: Kyunghyun Cho

10:30–11:00 Coffee break

**Session: Semantics and discourse** **Chair: Mats Wirén**

**11:00–12:30** Tagging Named Entities in 19th Century and Modern Finnish Newspaper Material with a Finnish Semantic Tagger **Kimmo Kettunen and Laura Löfberg**

Machine Learning for Rhetorical Figure Detection: More Chiasmus with Less Annotation **Marie Dubremetz and Joakim Nivre**

Coreference Resolution for Swedish and German using Distant Supervision **Alexander Wallin and Pierre Nugues**

12:30–13:30 Lunch

**Session: Lexical semantics** **Chair: Sara Stymne**

**13:30–14:30** Linear Ensembles of Word Embedding Models **Avo Muromägi, Kairit Sirts and Sven Laur**

Using Pseudowords for Algorithm Comparison: An Evaluation Framework for Graph-based Word Sense Induction **Flavio Massimiliano Cecchini, Chris Biemann and Martin Riedl**

14:30–15:00 Lightning talks

15:00–15:30 Coffee break

15:30–16:30 Poster session

16:30–17:15 Tutorial on Swedish: Maia Andréasson

17:15–18:00 SIG-ICALL meeting

19:00– Conference dinner

## Tuesday 23 May: Main Conference

| Room: Antarctic                     |  |
|-------------------------------------|--|
| 08:30–                              | Registration   |
| 09:00–09:30                         | Opening  |
| 09:30–10:30                         | Keynote: Kyunghyun Cho   |
| 10:30–11:00                         | Coffee break   |
| <b>Session: Mixed topics</b>        |  |
| <b>Chair: Trond Trosterud</b>       |  |
| <b>11:00–12:30</b>                  | <b>Aligning Phonemes Using Finite-state Methods</b>  |
|                                     | <b>Kimmo Koskenniemi</b>   |
|                                     | <b>Acoustic Model Compression with MAP Adaptation (student paper)</b>  |
|                                     | <b>Katri Leino and Mikko Kurimo</b>  |
|                                     | <b>OCR and Post-correction of Historical Finnish Texts (student paper)</b>   |
|                                     | <b>Senka Drobac, Pekka Kauppinen and Krister Lindén</b>  |
| <b>12:30–13:30</b> Lunch            |  |
| <b>Session: Machine translation</b> |  |
| <b>Chair: Aarne Ranta</b>           |  |
| <b>13:30–14:30</b>                  | <b>North-Sámi to Finnish Rule-based Machine Translation System</b>   |
|                                     | <b>Tommi Pirinen, Francis M. Tyers, Trond Trosterud, Ryan Johnson, Kevin Unhammer and Tiina Puolakainen</b>                                    |
|                                     | <b>Machine Translation with North Saami as a Pivot Language</b>  |
|                                     | <b>Lene Antonsen, Ciprian Gerstenberger, Maja Kappfjell, Sandra Nystø Ráhka, Marja-Liisa Olthuis, Trond Trosterud and Francis Morton Tyers</b> |
| 14:30–15:00                         | Lightning talks  |
| 15:00–15:30                         | Coffee break   |
| 15:30–16:30                         | Poster session   |
| 16:30–17:15                         | Tutorial on Swedish: Maia Andréasson   |
| <b>17:15–18:00</b>                  | <b>SIG-ICALL meeting</b>   |
| 19:00–                              | Conference dinner  |

## Tuesday 23 May: Main Conference

| Poster and demo session  |   |
|--|---|
| Cross-lingual Learning of Semantic Textual Similarity with Multilingual Word Representations       | Johannes Bjerva and Robert Östling  |
| A Modernised Version of the Glossa Corpus Search System  | Anders Nøklestad, Kristin Hagen, Janne Bondi Johannessen, Michał Kosek and Joel Priestley |
| Docforia: A Multilayer Document Model  | Marcus Klang and Pierre Nugues  |
| Dep_search: Efficient Search Tool for Large Dependency Parsebanks                                  | Juhani Luotolahti, Jenna Kanerva and Filip Ginter   |
| Finnish Resources for Evaluating Language Model Semantics  | Viljami Venekoski and Jouko Vankka  |
| The Effect of Translationese on Tuning for Statistical Machine Translation                         | Sara Stymne   |
| Multilingwis <sup>2</sup> – Explore Your Parallel Corpus   | Johannes Graën, Dominique Sandoz and Martin Volk  |
| Will my Auxiliary Tagging Task Help? Estimating Auxiliary Tasks Effectivity in Multi-Task Learning | Johannes Bjerva   |
| Iconic Locations in Swedish Sign Language: Mapping Form to Meaning with Lexical Databases          | Carl Börstell and Robert Östling  |
| Málrómur: A Manually Verified Corpus of Recorded Icelandic Speech                                  | Jón Guðnason, Sigrún Helgadóttir and Eiríkur Rögnvaldsson                                 |
| Proto-Indo-European Lexicon: The Generative Etymological Dictionary of Indo-European Languages     | Jouna Pyysalo   |
| Tilde MODEL - Multilingual Open Data for EU Languages  | Roberts Rozis and Raivis Skadiņš  |

## Wednesday 24 May: Main Conference

### Room: Wallenbergsalen

|  |   |
|--|---|
| <b>09:00–10:00</b> Keynote: Sharon Goldwater   | <b>Chair: Lilja Øvrelid</b>   |
| <b>10:00–10:30</b> Lightning talks   | <b>Chair: Jörg Tiedemann</b>  |
| 10:30–11:30 Coffee break + poster and demo session   |   |
| <b>11:30–12:30</b> Business meeting  |   |
|  |   |
| 12:30–13:30 Lunch  |   |
|  |   |
| <b>Session: Syntax</b>   | <b>Chair: Joakim Nivre</b>  |
| <b>13:30–15:00</b> SWEGRAM – A Web-Based Tool for Automatic Annotation and Analysis of Swedish Texts | Jesper Näsman, Beáta Megyesi and Anne Palmér  |
| Optimizing a PoS Tagset for Norwegian Dependency Parsing   | Petter Hohle, Lilja Øvrelid and Erik Velldal  |
| Creating Register Sub-corpora for the Finnish Internet Parsebank                                     | Veronika Laippala, Juhani Luotolahti, Aki-Juhani Kyröläinen, Tapio Salakoski and Filip Ginter |
|  |   |
| 15:00–15:30 Coffee break   |   |
| <b>15:30–16:30</b> Keynote: Rada Michalcea   | <b>Chair: Beáta Megyesi</b>   |
| <b>16:30–16:45</b> Closing   |   |

## Wednesday 24 May: Main Conference

### Room: Europe

|  |  |
|--|--|
| 09:00–10:00 Keynote: Sharon Goldwater  |  |
| 10:00–10:30 Lightning talks  |  |
| 10:30–11:30 Coffee break + poster and demo session   |  |
| 11:30–12:30 Business meeting   |  |
|  |  |
| 12:30–13:30 Lunch  |  |
|  |  |
| <b>Session: NLP applications</b>   | <b>Chair: Koenraad de Smedt</b>  |
| <b>13:30–15:00</b> KILLE: a Framework for Situated Agents for Learning Language Through Interaction  | Simon Dobnik and Erik de Graaf   |
| Data Collection from Persons with Mild Forms of Cognitive Impairment and Healthy Controls – Infrastructure for Classification and Prediction of Dementia | Dimitrios Kokkinakis, Kristina Lundholm Fors, Eva Björkner and Arto Nordlund |
| Evaluation of Language Identification Methods Using 285 Languages  | Tommi Jauhiainen, Krister Lindén and Heidi Jauhiainen                        |
|  |  |
| 15:00–15:30 Coffee break   |  |
| 15:30–16:30 Keynote: Rada Mihalcea   |  |
| 16:30–16:45 Closing  |  |

## Wednesday 24 May: Main Conference

### Room: Antarctic

09:00–10:00 Keynote: Sharon Goldwater

10:00–10:30 Lightning talks

10:30–11:30 Coffee break + poster and demo session

11:30–12:30 Business meeting

12:30–13:30 Lunch

#### Semantics

Chair: Shafqat Virk

**13:30–14:30 Can We Create a Tool for General Domain Event Analysis?**

**Siim Orasmaa and Heiki-Jaan Kaalep**

**From Treebank to Propbank: A Semantic-role and VerbNet Corpus for Danish**

**Eckhard Bick**

15:00–15:30 Coffee break

15:30–16:30 Keynote: Rada Mihalcea

16:30–16:45 Closing

## Wednesday 24 May: Main Conference

### Poster and demo session

The Effect of Excluding Out of Domain Training Data from Supervised Named-Entity Recognition (student paper) Adam Persson

Improving Optical Character Recognition of Finnish Historical Newspapers with a Combination of Fraktur & Antiqua Models and Image Preprocessing (student paper) Miika Koistinen, Kimmo Kettunen and Tuula Pääkkönen

Mainstreaming August Strindberg with Text Normalization (student paper) Adam Ek and Sofia Knuutinen

Quote Extraction and Attribution from Norwegian Newspapers Andrew Salway, Paul Meurer, Knut Hofland and Øystein Reigem

Cross-Lingual Syntax: Relating Grammatical Framework with Universal Dependencies Aarne Ranta, Prasanth Kolachina and Thomas Hallgren

Wordnet Extension via Word Embeddings: Experiments on the Norwegian Wordnet Heidi Sand, Erik Velldal and Lilja Øvrelid

Redefining Context Windows for Word Embedding Models: An Experimental Study Pierre Lison and Andrey Kutuzov

Exploring Properties of Intralingual and Interlingual Association Measures Visually Johannes Graën and Christof Bless

Universal Dependencies for Swedish Sign Language Robert Östling, Carl Börstell, Moa Gärdenfors and Mats Wirén

Exploring Treebanks with INESS Search Victoria Rosén, Helge Dyvik, Paul Meurer and Koenraad De Smedt

Word Vectors, Reuse, and Replicability: Towards a Community Repository of Large-text Resources Murhaf Fares, Andrey Kutuzov, Stephan Oepen and Erik Velldal

TALERUM - Learning Danish by Doing Danish Peter Juel Henrichsen

Services for Text Simplification and Analysis Johan Falkenjack, Evelina Rennes, Daniel Fahlborg, Vida Johansson and Arne Jönsson

A System for Identifying and Exploring Text Repetition in Large Historical Document Corpora Alekski Vesanto, Asko Nivala, Tapio Salakoski, Hannu Salmi and Filip Ginter