

# Conference Program

The 20th Nordic Conference of Computational Linguistics

**NODALIDA 2015**

May 11-13, 2015  
Institute of the Lithuanian Language  
Vilnius, Lithuania

## **Information**

### **Venue**

Welcome to the Institute of the Lithuanian Language!



### **Registration**

You can register in the lobby of the main entrance of the institute. The registration desk is open:

May 11: 08:30-10:00, 12:30-14:00, 18:30-19:30

May 12: 08:30-10:00

You can also register at the welcome reception (May 11) if you won't be able to attend any workshop.

### **Lunch and coffee**

Lunches and coffee are served in the lobby on the first and second floor.

### **Welcome reception May 11, 2015 19:00**

At the Institute of the Lithuanian Language

### **Conference Dinner May 12, 2015 19:00**

At the Restorant "Terrazza"

Address: Konstitucijos pr. 26, FORUM PALACE

## **Monday, May 11, 2015 NODALIDA WORKSHOPS**

- 09:00-13:00            **Constraint Grammar - Methods, Tools and Applications**  
Library
- 09:30-17:30            **4th Workshop on NLP for Computer Assisted Language Learning**  
Room 227
- 13:00-17:30            **Semantic resources and semantic annotation for Natural  
Language Processing and the Digital Humanities**  
Exhibition
- 13:30-17:40            **Innovative Corpus Query and Visualization Tools**  
Library
- Coffee breaks: 10:30-11:00, 15:00-15.30  
Lunch: 12:15-13:30
- 19:00                    **Welcome reception**

## Tuesday, May 12, 2015 MAIN CONFERENCE

09:10 – 09:30                    **Opening**    Main Hall

09:30 – 10:30                   **Keynote**    Main Hall

*Kevin Knight*

**How Much Information Does a Human Translator Add to the Original?**

Chair: Beáta Megyesi

10:30-11:00                    Coffee Break

11:00-12:30                   **Regular papers**

Parallel Session 1 <i>Syntax</i> Main hall Chair: Lars Ahrenberg	Parallel Session 2 <i>Annotation, Lithuanian NLP</i> Room 227 Chair: Jolanta Zabarskaitė	Parallel Session 3 <i>Speech</i> Exhibition Chair: Krister Lindén
<i>Eckhard Bick and Tino Didriksen.</i> <b>CG-3 – Beyond Classical Constraint Grammar</b>	<i>Yvonne Adesam, Gerlof Bouma and Richard Johansson.</i> <b>Defining the Eukalyptus forest – the Koala treebank of Swedish</b>	<i>Robert Reynolds and Francis Tyers.</i> <b>Automatic Word Stress Annotation of Russian Unrestricted Text</b>
<i>Jostein Lien, Erik Velldal and Lilja Øvrelid.</i> <b>Improving Cross-Domain Dependency Parsing with Dependency-Derived Clusters</b>	<i>Loïc Boizou, Jolanta Kovalevskaitė and Erika Rimkutė.</i> <b>Automatic Lemmatisation of Lithuanian MWEs</b>	<i>Dominik Sacha, Yuki Asano, Christian Rohrdantz, Felix Hamborg, Daniel Keim, Bettina Braun and Miriam Butt.</i> <b>Self Organizing Maps for the Visual Analysis of Pitch Contours</b>
<i>Jörg Tiedemann.</i> <b>Improving the Cross-Lingual Projection of Syntactic Dependencies</b>	<i>Jurgita Kapočiūtė-Dzikienė, Ligita Šarkutė and Andrius Utkas.</i> <b>The Effect of Author Set Size in Authorship Attribution for Lithuanian</b>	<i>Julián Zapata and Andreas Søborg Kirkedal.</i> <b>Assessing the Performance of Automatic Speech Recognition Systems When Used by Native and Non-Native Speakers of Three Major Languages in Dictation Workflows</b>

12:30-13:30                    Lunch

13:30-14:30

**Lightning talks**

Parallel Session 1 <i>Syntax and Semantics</i> Main hall Chair: Bolette Sanford Pedersen	Parallel Session 2 <i>Sign language and Speech</i> Room 227 Chair: Lars Borin
<i>Johan Bos and Malvina Nissim.</i> <b>Uncovering Noun-Noun Compound Relations by Gamification</b>	<i>Robert Östling, Carl Börstell and Lars Wallin.</i> <b>Enriching the Swedish Sign Language Corpus with Part of Speech Tags Using Joint Bayesian Word Alignment and Annotation Transfer</b>
<i>Héctor Martínez Alonso, Anders Johannsen, Barbara Plank and Anders Søgaard.</i> <b>Active Learning for Sense Annotation</b>	<i>Peter Juel Henriksen.</i> <b>Talebob - an Interactive Speech Trainer for Danish</b>
<i>Ildikó Pilán.</i> <b>Helping Swedish Words Come to their Senses: Word-Sense Disambiguation Based on Sense Associations from the SALDO Lexicon</b>	<i>Askars Salimbajevs and Jevgenijs Strigins.</i> <b>Using Sub-Word N-Gram Models for Dealing with OOV in Large Vocabulary Speech Recognition for Latvian</b>

14:30-15:00

Coffee break

14:30-16:00

**Posters and demos**

**Posters**

Chair: Elena Volodina

Lobby

*Johan Bos and Malvina Nissim.*

**Uncovering Noun-Noun Compound Relations by Gamification**

*Peter Juel Henriksen.*

**Talebob - an Interactive Speech Trainer for Danish**

*Héctor Martínez Alonso, Anders Johannsen, Barbara Plank and Anders Søgaard.*

**Active Learning for Sense Annotation**

*Ildikó Pilán.*

**Helping Swedish words Come to their Senses: Word-Sense Disambiguation Based on Sense Associations from the SALDO Lexicon**

*Askars Salimbajevs and Jevgenijs Strigins.*

**Using Sub-Word N-Gram Models for Dealing with OOV in Large Vocabulary Speech Recognition for Latvian**

*Robert Östling, Carl Börstell and Lars Wallin.*

**Enriching the Swedish Sign Language Corpus with Part of Speech Tags Using Joint Bayesian Word Alignment and Annotation Transfer**

**Demos**

Demo room

Chair: Violeta Meiliūnaitė

*Magnus Birkenes, Lars Johnsen, Arne Lindstad and Johanne Ostad.***From Digital Library to N-Grams: NB N-gram***Janne Bondi Johannessen.***The Corpus of American Norwegian Speech (CANS)***Johan Bos.***Open-Domain Semantic Parsing with Boxer***Sam Hardwick, Miikka Silfverberg and Krister Lindén.***Extracting Semantic Frames using hfst-pmatch***Arne Neumann.***discoursegraphs: A Graph-Based Merging Tool and Converter for Multilayer Annotated Corpora***Tommi A Pirinen.***Omorfi-Free and Open Source Morphological Lexical Database of Finnish***Evelina Rennes and Arne Jönsson.***A Tool for Automatic Simplification of Swedish Texts**

16:00-17:00

**Keynote**

Main Hall

*Catherine Pelachaud***Modeling socio-emotional humanoid agent**

Chair: Beáta Megyesi

19:00

Conference Dinner

Restaurant Terrazza

## Wednesday, May 13, 2015 MAIN CONFERENCE

09:00 -10:00                      **Keynote**    Main hall  
*Sebastian Riedel*  
**Embedding Probabilistic Logic for Machine Reading**  
 Chair: Filip Ginter

10:00 -10:30                      Coffee break

10:30-12:00                      **Regular and student papers**

Parallel Session 1 <i>Semantics</i> Main hall Chair: Jörg Tiedemann	Parallel Session 2 <i>Syntax and Semantics</i> Exhibition Chair: Lilja Øvrelid	Parallel Session 3 <i>Student Session</i> Room 227 Chair: Sofie Johansson Kokkinakis
<i>Juhani Luotolahti and Filip Ginter.</i> <b>Sentence Compression For Automatic Subtitling</b>	<i>Malin Ahlberg, Peter Andersson, Markus Forsberg and Nina Tahmasebi.</i> <b>A Case Study on Supervised Classification of Swedish Pseudo-Coordination</b>	<i>Max Berggren, Jussi Karlgren, Robert Östling and Mikael Parkvall.</i> <b>Inferring the Location of Authors from Words in their Texts</b>
<i>Sigrid Klerke, Héctor Martínez Alonso and Anders Søgaard.</i> <b>Looking hard: Eye Tracking for Detecting Grammaticality of Automatically Compressed Sentences</b>	<i>Eva Pettersson and Joakim Nivre.</i> <b>Improving Verb Phrase Extraction from Historical Text by use of Verb Valency Frames</b>	<i>Inari Listenmaa and Francis M. Tyers.</i> <b>Automatic Conversion of Colloquial Finnish to Standard Finnish</b>
<i>Richard Johansson and Luis Nieto Piña.</i> <b>Combining Relational and Distributional Knowledge for Word Sense Disambiguation</b>	<i>Sampo Pyysalo, Jenna Kanerva, Anna Missilä, Veronika Laippala and Filip Ginter.</i> <b>Universal Dependencies for Finnish</b>	<i>Scharolta Sienčnik.</i> <b>Adapting word2vec to Named Entity Recognition</b>

12:00-13:00                      Lunch

13:00–14:00                      **Business meeting**

14:00-14.30                      Coffee break

14:00-15:30  
Chair: Andrius Utka

**Poster session**

Lobby

Long papers:

*Héctor Martínez Alonso, Anders Johannsen, Sussi Olsen, Sanni Nimb, Nicolai Hartvig Sørensen, Anna Braasch, Anders Sjøgaard and Bolette Sandford Pedersen.*

**Supersense Tagging for Danish**

*Alexandr Chernov, Volha Petukhova and Dietrich Klakow.*

**Linguistically Motivated Question Classification**

*Jana Götze and Johan Boye.*

**Resolving Spatial References using Crowdsourced Geographical Data**

*Veronika Laippala, Jenna Kanerva, Anna Missilä, Sampo Pyysalo, Tapio Salakoski and Filip Ginter.*

**Towards the Classification of the Finnish Internet Parsebank: Detecting Translations and Informality**

*Mihai Lintean and Vasile Rus.*

**An Optimal Quadratic Approach to Monolingual Paraphrase Alignment**

*Michael Nokel and Natalia Loukachevitch.*

**Topic Models: Accounting Component Structure of Bigrams**

Short papers:

*Katarina Heimann Mühlenbock, Sofie Johansson Kokkinakis, Caroline Liberg, Åsa af Geijerstam, Jenny Folkeryd, Arne Jönsson, Erik Kanebrant and Johan Falkenjack.*

**A Multivariate Model for Classifying Texts' Readability**

*Gustavo Henrique Paetzold.*

**Using Positional Suffix Trees to Perform Agile Tree Kernel Calculation**

*Steinþór Steingrímsson, Sigrún Helgadóttir and Eiríkur Rögnvaldsson.*

**Analysing Inconsistencies and Errors in PoS Tagging in two Icelandic Gold Standards**

Student papers:

*Vytautas Mickevičius, Tomas Krilavičius Vaidas Morkevičius and Aušra Mackutė-Varoneckienė.*

**Automatic Thematic Classification of the Titles of the Seimas Votes**

*Birgitta Ojamaa, Päivi Kristiina Jokinen and Kadri Muischenk.*

**Sentiment Analysis on Conversational Texts**

15:30-16:30

**Tutorial on Lithuanian**

Main hall

*Algirdas Saugdargas.*

**Lithuanian Language in the Architecture of Knowledge**

Chair: Jolanta Zabarskaitė

16:30-16:40

**Closing**



## Sponsors

