

# Conference program

NODALIDA-2011

## 11 May

### Satellite events

#### Workshops

- Workshop on Creation, Harmonization and Application of Terminology Resources
- Workshop in Constraint Grammar Applications
- Workshop on Visibility and Availability of LT resources

19.00            *Welcome reception*

## 12 May

9.00–9.30        **Opening**

Mārcis Auziņš (Rector of the University of Latvia)  
Janne Bondi Johannessen (President of NEALT)  
Inguna Skadiņa (Chair of the Local Organizing Committee)  
Bolette Sandford Pedersen (Chair of the Program Committee)

9.30–10.30      **Invited Talk** (Chair: Costanza Navarretta)

Prof. Bonnie Webber (University of Edinburgh). *Discourse Structures and Language Technologies*

10.30–11.00     *Coffee*

11.00–13.00      **3 parallel sessions: REGULAR papers**

<b>Corpus creation, annotation and use</b> (Chair: Eiríkur Rögnvaldsson)	
11.00–11.30	Costanza Navarretta, Elisabeth Ahlsén, Jens Allwood, Kristiina Jokinen and Patrizia Paggio. <i>Creating Comparable Multimodal Corpora for Nordic Languages</i>
11.30–12.00	Rico Sennrich and Martin Volk. <i>Iterative, MT-based Sentence Alignment of Parallel Texts</i>
12.00–12.30	Estelle Delpech. <i>Evaluation of Terminologies Acquired from Comparable Corpora: an Application Perspective</i>
12.30–13.00	Janne Bondi Johannessen and Emiliano Raúl Guevara. <i>What Kind of Corpus is a Web Corpus?</i>
<b>Text and language classification</b> (Chair: Hanne Fersøe)	
11.00–11.30	Taraka Rama and Lars Borin. <i>Estimating Language Relationships from a Parallel Corpus. A Study of the Europarl Corpus</i>
11.30–12.00	Robert Remus. <i>Improving Sentence-level Subjectivity Classification through Readability Measurement</i>
12.00–12.30	Michael Wiegand and Dietrich Klakow. <i>Convolution Kernels for Subjectivity Detection</i>
<b>Morphology and POS tagging</b> (Chair: Janne Bondi Johannessen)	
11.00–11.30	Miikka Silfverberg and Krister Lindén. <i>Combining Statistical Models for POS Tagging using Finite-State Calculus</i>
11.30–12.00	Niels Beuck, Arne Köhn and Wolfgang Menzel. <i>Decision Strategies for Incremental POS Tagging</i>
12.00–12.30	Anssi Yli-Jyrä. <i>Explorations on Positionwise Flag Diacritics in Finite-State Morphology</i>
12.30–13.00	Heiki-Jaan Kaalep and Kadri Muischnek. <i>Morphological Analysis of a Non-Standard Language Variety</i>

13.00–14.00      *Lunch*

14.00–15.30 **12 Posters and Demos** (Chair: Anders Nøklestad)

<b>Wordnets and lexical issues</b>
<p>Kristiina Muhonen and Krister Lindén. <i>Do Wordnets also Improve Human Performance on NLP Tasks?</i></p> <p>Loïc Boizou. <i>The Formal Patterns of the Lithuanian Verb Forms</i></p> <p>Hector Martinez Alonso, Núria Bel and Bolette Sandford Pedersen. <i>Identification of Sense Selection in Regular Polysemy using Shallow Features</i></p> <p>Anders Johannsen and Bolette Sandford Pedersen. <i>“Andre ord” — a Wordnet Browser for the Danish Wordnet, DanNet (DEMO)</i></p>
<b>Syntax</b>
<p>Anne Bjerre. <i>Extraction from Relative and Embedded Interrogative Clauses in Danish</i></p> <p>Martin Hassel, Aron Henriksson and Sumithra Velupillai. <i>Something Old, Something New — Applying a Pre-trained Parsing Model to Clinical Swedish</i></p> <p>Atro Voutilainen and Tanja Purtonen. <i>A Double-blind Experiment on Interannotator Agreement: the Case of Dependency Syntax and Finnish</i></p> <p>Lauma Pretkalniņa, Gunta Nešpore, Kristīne Levāne-Petrova and Baiba Saulīte. <i>A Prague Markup Language Profile for the SemTi-Kamols Grammar Model</i></p> <p>Daiga Dekšne and Raivis Skadiņš. <i>CFG Based Grammar Checker for Latvian</i></p>
<b>Morphology</b>
<p>Sami Virpioja, Oskar Kohonen and Krista Lagus. <i>Evaluating the Effect of word Frequencies in a Probabilistic Generative Model of Morphology</i></p> <p>Tommi Pirinen. <i>Modularisation of Finnish Finite-State Language Description — Towards Wide Collaboration in Open Source Development of a Morphological Analyser</i></p>
<b>Machine translation</b>
<p>Sara Stymne. <i>Iterative Reordering and Word Alignment for Statistical MT</i></p>

15.30–15.45 *Coffee*

15.45–17.15      **3 parallel sessions: REGULAR papers**

<b>Speech</b> (Chair: Meelis Mihkla)	
15.45–16.15	Sofia Strömbergsson. <i>Corrective Re-synthesis of Deviant Speech Using Unit Selection</i>
16.15–16.45	Peter Juel Henriksen. <i>Fishing in a Speech Stream, Angling for a Lexicon</i>
16.45–17.15	Bea Valkenier, Dirkjan Krijnders, Ronald van Elburg and Tjeerd Andringa. <i>Psycho-Acoustically Motivated Formant Feature Extraction</i>
<b>Search and information extraction</b> (Chair: Costanza Navarretta)	
15.45–16.15	Lars Borin, Markus Forsberg and Christer Ahlberger. <i>Semantic Search in Literature as an e-Humanities Research Tool: CONPLISIT — Consumption Patterns and Life-Style in 19th Century Swedish Literature</i>
16.15–16.45	Silja Huttunen, Arto Vihavainen and Roman Yangarber. <i>Relevance Prediction in Information Extraction Using Discourse and Lexical Features</i>
16.45–17.15	Gintarė Grigonytė, Erika Rimkutė, Andrius Utkā and Loïc Boizou. <i>Experiments on Lithuanian Term Extraction</i>
<b>Syntax, indexing</b> (Chair: Jussi Karlgren)	
15.45–16.15	Peter Ljunglöf. <i>Editing Syntax Trees on the Surface</i>
16.15–16.45	Anders Søgaard. <i>Using Graphical Models for PP Attachment</i>
16.45–17.15	Erik Velldal. <i>Random Indexing Re-Hashed</i>

17.15–18.15      **Invited Talk** (Chair: Inguna Skadiņa)

Prof. Guntis Bārzdīņš (University of Latvia). *When FrameNet Meets a Controlled Natural Language*

19.30              **Conference dinner**

## 13 May

9.00–10.00      **Invited Talk** (Chair: Kristiina Jokinen)  
Prof. Joakim Nivre (Uppsala University). *Bare-Bones Dependency Parsing — A Case for Occam's Razor?*

10.00–10.30      *Coffee*

10.30–12.00      **3 parallel sessions: REGULAR papers and STUDENT papers**

<b>Lexicon, etymology</b> (Chair: Bolette Sandford Pedersen)	
10.30–11.00	Eckhard Bick. <i>A FrameNet for Danish</i>
11.00–11.30	Ingemar Hjälmsstad, Martin Hassel and Maria Skeppstedt. <i>The Impact of Part-of-Speech Filtering on Generation of a Swedish-Japanese Dictionary using English as Pivot Language</i>
11.30–12.00	Hannes Wettig and Roman Yangarber. <i>Probabilistic Models for Alignment of Etymological Data</i>
<b>Machine translation; classification</b> (Chair: Andrejs Vasiļjevs)	
10.30–11.00	Martin Volk and Rico Sennrich. <i>Disambiguation of English Contractions for Machine Translation of TV Subtitles</i>
11.00–11.30	Raivis Skadiņš, Tatiana Gornostay and Valters Šics. <i>Toponym Disambiguation in English-Lithuanian SMT System with Spatial Knowledge</i>
11.30–12.00	Eirini Florou and Stasinou Konstantopoulos. <i>A Quantitative and Qualitative Analysis of Nordic Surnames</i>
<b>Student papers</b> (Chair: Normunds Grūzītis)	
10.30–11.00	Marius Olausson. <i>Evaluating the Speech Quality of the Norwegian Synthetic Voice Brage</i>
11.00–11.30	Mojgan Seraji. <i>A Statistical Part-of-Speech Tagger for Persian</i>
11.30–12.00	Angel Genov and Georgi Iliev. <i>Linguistic Motivation in Automatic Sentence Alignment of Parallel Corpora: the Case of Danish-Bulgarian and English-Bulgarian</i>

12.00–13.00      *Lunch*

13.00–14.30      **12 Posters/demos** (Chair: Kristiina Jokinen)

<b>Classification &amp; summarization</b>
<p>Christian Smith and Arne Jönsson. <i>Automatic Summarization as Means of Simplifying Texts, an Evaluation for Swedish</i></p> <p>Per Almquist and Jussi Karlgren. <i>Experiments to Investigate the Utility of Nearest Neighbor Metrics Based on Linguistically Informed Features for Detecting Textual Plagiarism</i></p> <p>Jody Foo. <i>A Categorization Scheme for Analyzing Rules from a Handbook of Swedish Writing Rules</i></p> <p>Anju Saxena and Lars Borin. <i>Dialect Classification in the Himalayas: a Computational Approach</i></p>
<b>Knowledge systems</b>
<p>Ann-Marie Eklund. <i>Query Constraining Aspects of Knowledge</i></p> <p>Kenneth Wilhelmsson. <i>Automatic Question Generation from Swedish Documents as a Tool for Information Extraction</i></p>
<b>Corpus creation, annotation and use</b>
<p>Christian Hänig. <i>Knowledge-free Verb Detection through Sentence Sequence Alignment</i></p> <p>Maria Holmqvist and Lars Ahrenberg. <i>A Gold Standard for English-Swedish Word Alignment</i></p> <p>Anne-Kathrin Schumann. <i>Corpus-based Terminology: Detection, Description and Representation of Knowledge-rich Contexts in Russian</i></p>
<b>Student posters</b>
<p>Anastasia Shimorina and Maria Grachkova. <i>Identification of Context Markers for Russian Nouns</i></p> <p>Gideon Kotzé. <i>Finding Statistically Motivated Features Influencing Subtree Alignment Performance</i></p>

14.30–15.30      **NEALT Business meeting**

15.30–16.00      **Closing**

16.00-16.30      *Coffee*