

Conference program NODALIDA-2007 Main Conference

Friday, May 25, 2007

9.00–10.30 **Opening Session**

9.00– 9.30 Opening

9.30–10.30 Invited Talk

Evaluating Automatic Approaches for Word Meaning Discovery and Disambiguation using Lexical Substitution

Diana F. McCarthy, University of Sussex

10.30–11.00 **Coffee Break**

11.00–12.30 **Parallel Paper Sessions**

| | Parsing | Multilingual Resources and Translation | Speech Technology |
|-------------|--|--|--|
| 11.00–11.30 | <i>IceParser: An Incremental Finite-State Parser for Icelandic</i> Hrafn Loftsson and Eiríkur Rögnvaldsson | <i>Using Danish as a CG Interlingua: A Wide-Coverage Norwegian-English Machine Translation System</i> Eckhard Bick and Lars Nygaard | <i>Automatic Compound Word Reconstruction for Speech Recognition of Compounding Languages</i> Tanel Alumäe |
| 11.30–12.00 | <i>Combining Contexts in Lexicon Learning for Semantic Parsing</i> Richard Socher, Chris Biemann and Rainer Osswald | <i>Identifying Cross Language Term Equivalents Using Statistical Machine Translation and Distributional Association Measures</i> Hans Hjelm | <i>Multivariate Cepstral Feature Compensation on Band-limited Data for Robust Speech Recognition</i> Nicolas Morales, John H. L. Hansen, Doroteo T. Toledano and Javier Garrido |
| 12.00–12.30 | <i>Polynomial Charts for Totally Unordered Languages</i> Anders Søgaard | <i>The Swedish-Turkish Parallel Corpus and Tools for its Creation</i> Beata B. Megyesi and Bengt Dahlqvist | <i>Development of Text-To-Speech System for Latvian</i> Kārlis Goba and Andrejs Vasiļjevs |

12.30–14.00 **Lunch**

14.00–15.30 **Parallel Paper Sessions**

| | Treebanks | Information Extraction and Summarization | Phonetics and Speech |
|-------------|--|--|--|
| 14.00–14.30 | <i>Theoretically Motivated Treebank Coverage</i> Victoria Rosén and Koenraad De Smedt | <i>Time Extraction from Real-time Generated Football Reports</i> Markus Borg | <i>Utterance-initial Duration of Finnish Non-plosive Consonants</i> Tuomo Saarni, Jussi Hakokari, Olli Aaltonen, Jouni Isoaho and Tapio Salakoski |
| 14.30–15.00 | <i>Extended Constituent-to-Dependency Conversion for English</i> Richard Johansson and Pierre Nugues | <i>The Extraction of Trajectories from Real Texts Based on Linear Classification</i> Hanjing Li, Tiejun Zhao, Sheng Li, Jiyuan Zhao | <i>An Advanced Speech Corpus for Norwegian</i> Janne Bondi Johannessen, Kristin Hagen, Joel James Priestley and Lars Nygaard |
| 15.00–15.30 | <i>Clausal Coordinate Ellipsis in German: The TIGER Treebank as a Source of Evidence</i> Karin Harbusch and Gerard Kempen | <i>Widening the HolSum Search Scope</i> Martin Hassel and Jonas Sjöbergh | <i>Spoken Document Retrieval in a Highly Inflectional Language</i> Inger Ekman and Kalervo Järvelin |

15.30–16.00 **Coffee Break**

16.00–17.00 **Poster Session 1**

LinES: An English-Swedish Parallel Treebank

Lars Ahrenberg

Posterior Probability Based Confidence Measures Applied to a Children's Speech Reading Tracking System

Daniel Bolanos and Wayne H. Ward

Estonian-English Statistical Machine Translation: the First Results

Mark Fishel, Heiki-Jaan Kaalep and Kadri Muischnek

A Hybrid Constituency-Dependency Parser for Swedish

Johan Hall, Joakim Nivre and Jens Nilsson

Íslenskur Orðasjóður – Building a Large Icelandic Corpus

Erla Hallsteinsdóttir, Thomas Eckart, Chris Biemann, Uwe Quasthoff and Matthias Richter

A Survey and Classification of Methods for (Mostly) Unsupervised Learning

Harald Hammarström

Marvina – A Norwegian Speech-Centric, Multimodal Visitor Guide

Ole Hartvigsen, Erik Harborg, Tore Amble and Magne H. Johnsen

A Norwegian Letter-to-Sound Engine with Danish as a Catalyst

Peter Juel Henriksen

Dialogue Simulation and Context Dynamics for Dialogue Management

Simon Keizer and Roser Morante

Managing Keyword Variation with Frequency Based Generation of Word Forms in IR

Kimmo Kettunen

Developing and Evaluating a Searchable Swedish-Thai Lexicon

Wanwisa Khanaraksombat and Jonas Sjöbergh

Identification of Entity References in Hospital Discharge Letters

Dimitrios Kokkinakis and Anders Thurin

Lexical Parameters, Based on Corpus Analysis of English and Swedish Cancer Data, of Relevance for NLG

Dimitrios Kokkinakis, Maria Toporowska Gronostaj, Catalina Hallett and David Hardcastle

17.00–18.00 **Poster Session 2**

Anatomy of an XML-based Text Corpus Server

Mikko Lounela

Perceptual Assessment of the Degree of Russian Accent

Lya Meister

Terminology Extraction and Term Ranking for Standardizing Term Banks

Magnus Merkel and Jody Foo

Representing Calendar Expressions with Finite-State Transducers that Bracket Periods of Time on a Hierarchical Timeline

Jyrki Niemi and Kimmo Koskenniemi

Parsing Manually Detected and Normalized Disfluencies in Spoken Estonian

Helen Nigol

Designing a Speech Corpus for Estonian Unit Selection Synthesis

Liisi Piits, Meelis Mihkla, Tõnis Nurk and Indrek Kiissel

Evaluating Evaluation Measures

Ines Rehbein and Josef van Genabith

Development of a Modern Greek Broadcast-News Corpus and Speech Recognition System

Jürgen Riedler and Sergios Katsikas

Role of Different Spectral Attributes in Vowel Categorization: the Case of Udmurt

Janne Savela, Stina Ojala, Olli Aaltonen and Tapio Salakoski

Recreating Humorous Split Compound Errors in Swedish by Using Grammaticality

Jonas Sjöbergh and Kenji Araki

A Re-examination of Question Classification

Håkan Sundblad

Interpretation of Yes/No Questions as Metaphor Recognition

Tarmo Truu, Haldur Öim and Mare Koit

Rule-based Logical Forms Extraction

Cenny Wenner

19.00– Conference Dinner

Saturday, May 26, 2007

9.00–10.00 **Plenary Session**

Invited Talk

Text Analysis and Machine Learning for Stylometrics and Stylogenetics

Walter Daelemans, University of Antwerp

10.00–10.30 **Coffee Break**

10.30–12.30 **Parallel Student Sessions**

| | Spoken Language Processing | Multilingual Resources and Translation | Natural Language Processing |
|-------------|--|--|--|
| 10.30–11.00 | <i>Clause Boundary Detection in Transcribed Spoken Language</i> Fredrik Jørgensen | <i>Memory-Based Learning of Word Translation</i> Maria Holmqvist | <i>Unmediated Data-Oriented Generation</i> Dave Cochran |
| 11.00–11.30 | <i>The Effects of Disfluency Detection in Parsing Spoken Language</i> Fredrik Jørgensen | <i>Using Parallel Corpora to Create a Greek-English Dictionary with UPLUG</i> Konstantinos Charitakis | <i>Linguistically Fuelled Text Similarity</i> Björn Andrist and Martin Hassel |

| | | | |
|-------------|---|--|---|
| 11.30–12.00 | <i>Tagging a Norwegian Speech Corpus</i> Anders Nøklestad and Åshild Søfteland | <i>Using a Wizard of Oz as a Baseline to Determine Which System Architecture Is the Best for a Spoken Language Translation System</i> Marianne Starlander | <i>Decomposing Swedish Compounds Using Memory-Based Learning</i> Karin Friberg |
| 12.00–12.30 | <i>Initial Experiments with Estonian Speech Recognition</i> Anton Ragni | <i>Grammar Sharing Techniques for Rule-based Multilingual NLP Systems</i> Marianne Santaholma | <i>A Method for Resolution of Temporal Expressions in Estonian Natural Language Dialogue Systems</i> Margus Treumuth |

12.30–14.00 **Lunch**

14.00–15.30 **Parallel Paper Sessions**

| | Parsing and Translation | Machine Learning | Dialogue |
|-------------|---|--|---|
| 14.00–14.30 | <i>Dependency-Based Hybrid Model of Syntactic Analysis for the Languages with a Rather Free Word Order</i> Guntis Bārzdīns, Normunds Grūzītis, Gunta Nešpore and Baiba Saulīte | <i>Comparison of the Self-Organizing Map and Multidimensional Scaling in Analysis of Estonian Emotion Concepts</i> Toomas Kirt and Ene Vainik | <i>Interview and Delivery: Dialogue Strategies for Conversational Recommender Systems</i> Pontus Wärnestål, Lars Degerstedt and Arne Jönsson |
| 14.30–15.00 | <i>Comparing French PP-attachment to English, German and Swedish</i> Martin Volk and Frida Tidström | <i>Evaluating Stages of Development in Second Language French: A Machine-Learning Approach</i> Jonas Granfeldt and Pierre Nugues | <i>Achieving Goals in Collaboration: Analysis of Estonian Institutional Calls</i> Olga Gerassimenko, Mare Koit, Andriela Rääbis and Krista Strandson |

| | | | |
|-----------------|--|---|--|
| 15.00– 15.30 | <i>Comprehension Assistant for Languages of Baltic States</i> Inguna Skadiņa, Andrejs Vasiļjevs, Daiga Deksnē, Raivis Skadiņš and Linda Goldberga | <i>Inducing Baseform Models from a Swedish Vocabulary Pool</i> Eva Forsbom | |
|-----------------|--|---|--|

15.30–16.00 **Coffee Break**

16.00–17.30 **Closing Session**

16.00–17.00 Business Meeting of the Northern European Association for Language Technology (NEALT)

17.00–17.30 Closing