

# Contents

<b>Contents</b>	<b>iii</b>
<b>Preface</b>	<b>vii</b>
<b>Committees</b>	<b>ix</b>
<b>Conference Program</b>	<b>xi</b>
<b>I Invited Papers</b>	<b>1</b>
<b>JEAN CARLETTA</b> <b>Developing Meeting Support Technologies: From Data to Demonstration (and Beyond)</b>	<b>2</b>
<b>RALF STEINBERGER</b> <b>Linking News Content Across Languages</b>	<b>4</b>
<b>II Tutorial</b>	<b>6</b>
<b>GRAHAM WILCOCK</b> <b>Text Annotation with OpenNLP and UIMA</b>	<b>7</b>
<b>III Regular papers</b>	<b>9</b>
<b>LENE ANTONSEN, SAARA HUHMARNIEMI AND TROND TROSTERUD</b> <b>Interactive pedagogical programs based on constraint grammar</b>	<b>10</b>
<b>JARI BJÖRNE, FILIP GINTER, JUHO HEIMONEN, SAMPO PYYSALO AND TAPIO SALAKOSKI</b> <b>Learning to Extract Biological Event and Relation Graphs</b>	<b>18</b>
<b>HERCULES DALIANIS, MARTIN RIMKA AND VIGGO KANN</b> <b>Using Uplug and SiteSeeker to construct a cross language search engine for Scandinavian languages</b>	<b>26</b>
<b>EVA FORSBOM</b> <b>Extending the View: Explorations in Bootstrapping a Swedish PoS Tagger</b>	<b>34</b>
<b>TATIANA GORNOSTAY AND INGUNA SKADIŅA</b> <b>Pattern-based English-Latvian Toponym Translation</b>	<b>41</b>
<b>NATHAN GREEN, PAUL BREIMYER, VINAY KUMAR AND NAGIZA F SAMATOVA</b> <b>WebBANC: Building Semantically-Rich Annotated Corpora from Web User Annotations of Minority Languages</b>	<b>48</b>
<b>CHRISTIAN HARDMEIER AND MARTIN VOLK</b> <b>Using Linguistic Annotations in Statistical Machine Translation of Film Subtitles</b>	<b>57</b>

<b>KATRI HAVERINEN, FILIP GINTER, VERONIKA LAIPPALA AND TAPIO SALAKOSKI</b> <b>Parsing Clinical Finnish: Experiments with Rule-Based and Statistical Dependency Parsers</b>	<b>65</b>
<b>JANNE BONDI JOHANNESSEN, JOEL PRIESTLEY, KRISTIN HAGEN, TOR ANDERS ÅFARLI AND ØYSTEIN ALEXANDER VANGSNES</b> <b>The Nordic Dialect Corpus — an advanced research tool</b>	<b>73</b>
<b>PETER KOLB</b> <b>Experiments on the difference between semantic similarity and relatedness</b>	<b>81</b>
<b>KRISTER LINDÉN AND TOMMI PIRINEN</b> <b>Weighted Finite-State Morphological Analysis of Finnish Compounding with HFST-LEXC</b>	<b>89</b>
<b>KRISTER LINDÉN AND JUSSI TUOVILA</b> <b>Corpus-based Paradigm Selection for Morphological Entries</b>	<b>96</b>
<b>HRAFN LOFTSSON, IDA KRAMARCZYK, SIGRÚN HELGADÓTTIR AND EIRÍKUR RÖGNVALDSSON</b> <b>Improving the PoS tagging accuracy of Icelandic text</b>	<b>103</b>
<b>OLGA LASHEVSKAJA AND OLGA MITROFANOVA</b> <b>Disambiguation of Taxonomy Markers in Context: Russian Nouns</b>	<b>111</b>
<b>YVES LEPAGE AND CHOOI LING GOH</b> <b>Towards automatic acquisition of linguistic features</b>	<b>118</b>
<b>MIGUEL A. MOLINERO, BENOÎT SAGOT AND LIONEL NICOLAS</b> <b>Building a morphological and syntactic lexicon by merging various linguistic resources</b>	<b>126</b>
<b>KRISTINA NILSSON AND HANS HJELM</b> <b>Using Semantic Features Derived from Word-Space Models for Swedish Coreference Resolution</b>	<b>134</b>
<b>JACOB PERSSON, RICHARD JOHANSSON AND PIERRE NUGUES</b> <b>Text Categorization Using Predicate–Argument Structures</b>	<b>142</b>
<b>MAGNUS ROSELL</b> <b>Part of Speech Tagging for Text Clustering in Swedish</b>	<b>150</b>
<b>BOLETTE SANDFORD PEDERSEN AND ANNA BRAASCH</b> <b>What do we need to know about humans? A view into the DanNet database</b>	<b>158</b>
<b>NATALIE SCHLUTER AND JOSEF VAN GENABITH</b> <b>Dependency Parsing Resources for French: Converting Acquired Lexical Functional Grammar F-Structure Annotations and Parsing F-Structures Directly</b>	<b>166</b>
<b>MIIKKA SILFVERBERG AND KRISTER LINDÉN</b> <b>Conflict Resolution Using Weighted Rules in HFST-TWOLC</b>	<b>174</b>
<b>ANDERS SØGAARD</b> <b>A linear time extension of deterministic pushdown automata</b>	<b>182</b>
<b>ANDERS SØGAARD</b> <b>Verifying context-sensitive treebanks and heuristic parses in polynomial time</b>	<b>190</b>
<b>MICHAEL WIEGAND AND DIETRICH KLAKOW</b> <b>Predictive Features in Semi-Supervised Learning for Polarity Classification and the Role of Adjectives</b>	<b>198</b>

<b>ANSSI YLI-JYRÄ</b> <b>An Efficient Double Complementation Algorithm for Superposition-Based Finite-State Morphology</b>	<b>206</b>
<b>IV Regular short paper</b>	<b>214</b>
<b>ECKHARD BICK AND M. PILAR VALVERDE IBÁÑEZ</b> <b>Automatic Semantic Role Annotation for Spanish</b>	<b>215</b>
<b>MARK FISHEL AND JOAKIM NIVRE</b> <b>Voting and Stacking in Data-Driven Dependency Parsing</b>	<b>219</b>
<b>KARIN FRIBERG HEPPIN</b> <b>MedEval Six Test Collections in One</b>	<b>223</b>
<b>RASHMI GANGADHARAI AH, RALF D. BROWN AND JAIME CARBONELL</b> <b>Active Learning in Example-Based Machine Translation</b>	<b>227</b>
<b>ANTON K. INGASON, SKÚLI B. JÓHANNSSON, EIRÍKUR RÖGNVALDSSON, HRAFN LOFTSSON AND SIGRÚN HELGADÓTTIR</b> <b>Context-Sensitive Spelling Correction and Rich Morphology</b>	<b>231</b>
<b>MANFRED KLENNER, ANGELA FAHRNI AND STEFANOS PETRAKIS</b> <b>PolArt: A Robust Tool for Sentiment Analysis</b>	<b>235</b>
<b>BEÁTA B. MEGYESI</b> <b>The Open Source Tagger HunPoS for Swedish</b>	<b>239</b>
<b>INGUNA SKADIŅA AND EDGARS BRĀLĪTIS</b> <b>English-Latvian SMT: knowledge or data?</b>	<b>242</b>
<b>LILJA ØVRELID</b> <b>Cross-lingual porting of distributional semantic classification</b>	<b>246</b>
<b>V Student papers</b>	<b>250</b>
<b>MARIA ESKEVICH</b> <b>Prominence detected by listeners for future speech synthesis application</b>	<b>251</b>
<b>OKKO RĀSĀNEN AND JORIS DRIESEN</b> <b>A comparison and combination of segmental and fixed-frame signal representations in NMF-based word recognition</b>	<b>255</b>
<b>BĀLINT SASS</b> <b>Verb Argument Browser for Danish</b>	<b>263</b>
<b>VI Demos</b>	<b>267</b>
<b>ECKHARD BICK</b> <b>DeepDict — A Graphical Corpus-based Dictionary of Word Relations</b>	<b>268</b>
<b>SANDRA DERBRING, PETER LJUNGLÖF AND MARIA OLSSON</b> <b>SubTTS: Light-weight automatic reading of subtitles</b>	<b>272</b>
<b>PETER LJUNGLÖF, STAFFAN LARSSON, KATARINA MÜHLENBOCK AND GUNILLA THUNBERG</b> <b>TRIK: A Talking and Drawing Robot for Children with Communication Disabilities</b>	<b>275</b>

<b>BODIL NISTRUP MADSEN AND HANNE ERDMAN THOMSEN</b> <b>CAOS — A tool for the Construction of Terminological Ontologies</b>	<b>279</b>
<b>ARNE MARTINUS LINDSTAD, ANDERS NØKLESTAD, JANNE BONDI JOHANNESSEN AND ØYS-</b> <b>TEIN A. VANGSNES</b> <b>The Nordic Dialect Database: Mapping Microsyntactic Variation in the Scandinavian Lan-</b> <b>guages</b>	<b>283</b>
<b>Author Index</b>	<b>287</b>