# Table of Contents

## Text

Knowledge Sources for Word Sense Disambiguation .......................... 1

*E. Agirre and D. Martinez*

The Current Status of the Prague Dependency Treebank .................... 11

*E. Hajičová, J. Hajič, B. Hladká, M. Holub, P. Pajas, V. Řezníčková, and P. Sgall*

Language Corpora: The Czech Case ........................................... 21

*F. Čermák*

A Hybrid Logic Formalization of Information Structure Sensitive Discourse Interpretation .............................................. 31

*G.-J.M. Kruijff and I. Kruijff-Korbayová*

The Possibilities of Automatic Detection/Correction of Errors in Tagged Corpora: A Pilot Study on a German Corpus ....................... 39

*K. Oliva*

Grammatical Agreement and Automatic Morphological Disambiguation of Inflectional Languages ................................................. 47

*V. Petkevič*

Modelling Semantic Association and Conceptual Inheritance for Semantic Analysis .................................................. 54

*P. Vaillant*

Influence of Conditional Independence Assumption on Verb Subcategorization Detection .................................................. 62

*K. Kermanidis, M. Maragoudakis, N. Fakotakis, and G. Kokkinakis*

Morphological Guesser of Czech Words ........................................ 70

*J. Hlaváčová*

Clustering Technique Based on Semantics for Definite Description Resolution .................................................. 76

*R. Muñoz and M. Palomar*

Pronoun Resolution in Spanish from Full Parsing .......................... 84

*M. Saiz-Noeda, M. Palomar, and L. Moreno*

A Method of Accurate Robust Parsing of Czech ............................. 92

*V. Kubaň and M. Plátek*
<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>A New Czech Morphological Analyser a jka</td>
<td>100</td>
</tr>
<tr>
<td>R. Sedláček and P. Smrž</td>
<td></td>
</tr>
<tr>
<td>Finding Semantically Related Words in Large Corpora</td>
<td>108</td>
</tr>
<tr>
<td>P. Smrž and P. Rychlá</td>
<td></td>
</tr>
<tr>
<td>Syntactic-Based Methods for Measuring Word Similarity</td>
<td>116</td>
</tr>
<tr>
<td>P. Gamallo, C. Gasperin, A. Agustini, and G.P. Lopes</td>
<td></td>
</tr>
<tr>
<td>The ATRACT Workbench:</td>
<td></td>
</tr>
<tr>
<td>Automatic Term Recognition and Clustering for Terms</td>
<td>126</td>
</tr>
<tr>
<td>H. Mima, S. Ananiadou, and G. Nenadič</td>
<td></td>
</tr>
<tr>
<td>Augmented Auditory Representation of e-Texts for Text-to-Speech Systems</td>
<td>134</td>
</tr>
<tr>
<td>G. Xydas and G. Kouroupetrogou</td>
<td></td>
</tr>
<tr>
<td>Enhancing the Valency Dictionary of Czech Verbs:</td>
<td>142</td>
</tr>
<tr>
<td>Tectogrammatical Annotation</td>
<td></td>
</tr>
<tr>
<td>H. Skoumatová, M. Straňáková-Lapatková, and Z. Žabokrský</td>
<td></td>
</tr>
<tr>
<td>An Interactive Graph Based Legal Information Retrieval System</td>
<td>150</td>
</tr>
<tr>
<td>P. Quaresma and I.P. Rodrigues</td>
<td></td>
</tr>
<tr>
<td>Text Segmentation into Paragraphs Based on Local Text Cohesion</td>
<td>158</td>
</tr>
<tr>
<td>I.A. Bolshakov and A. Gelbukh</td>
<td></td>
</tr>
<tr>
<td>Building a Digital Collection of Web-Pages:</td>
<td>167</td>
</tr>
<tr>
<td>Access and Filtering Information with Textual Expansion</td>
<td></td>
</tr>
<tr>
<td>O. Larouk</td>
<td></td>
</tr>
<tr>
<td>Three Approaches to Word Sense Disambiguation for Czech</td>
<td>174</td>
</tr>
<tr>
<td>R. Král</td>
<td></td>
</tr>
<tr>
<td>Method for WordNet Enrichment Using WSD</td>
<td>180</td>
</tr>
<tr>
<td>A. Montoyo, M. Palomar, and G. Rigau</td>
<td></td>
</tr>
<tr>
<td>Speech</td>
<td>-------</td>
</tr>
<tr>
<td>Human Speech Perception: Some Lessons from Automatic Speech Recognition</td>
<td>187</td>
</tr>
<tr>
<td>H. Heřmanský</td>
<td></td>
</tr>
<tr>
<td>Speech Technology in Reality – Applications, Their Challenges and Solutions</td>
<td>197</td>
</tr>
<tr>
<td>H. Mangold</td>
<td></td>
</tr>
<tr>
<td>Generative Phoneme-Threephone Model for ASR</td>
<td>201</td>
</tr>
<tr>
<td>T.K. Vintsiuk</td>
<td></td>
</tr>
<tr>
<td>Algebraic Models of Speech Segment Databases</td>
<td>208</td>
</tr>
<tr>
<td>I. Kopecěk</td>
<td></td>
</tr>
</tbody>
</table>
## Table of Contents

Bayesian Noise Compensation of Time Trajectories of Spectral Coefficients for Robust Speech Recognition ........................................ 214
   I. Potamitis, N. Fakotakis, and G. Kokkinakis

The Influence of a Filter Shape in Telephone-Based Recognition Module Using PLP Parameterization ........................................... 222
   J. Psutka, L. Müller, and J.V. Psutka

System for Speech Communication through Internet .......................... 229
   A. Machovikov and K. Stolyarov

Phoneme Based ASR System for Slovak SpeechDat Database ............. 237
   M. Marcinek, J. Juhar, and A. Cizmar

Large Vocabulary Continuous Speech Recognizer for Slovenian Language .... 242
   A. Žgank, Z. Kačič, and B. Horvat

A Nonlinearized Discriminant Analysis and Its Application to Speech Impediment Therapy .................................................. 249
   A. Kocsor, L. Tóth, and D. Paczolay

Speech Recognition Issues for Dutch Spoken Document Retrieval .......... 258
   R. Ordelman, A. van Hessen, and F. de Jong

Multichannel Sound Acquisition with Stress Situations Determination for Medical Supervision in a Smart House ............................... 266
   D. Istrate and E. Castelli

Two-Pass Recognition of Czech Speech Using Adaptive Vocabulary .......... 273
   P. Ircing and J. Psutka

Simulation of Speaking Styles with Adapted Prosody ........................ 278
   H. Kruschke

The Utterance Type and the Size of the Intonation Unit .................... 285
   N. Volskaya and P. Skrelin

Estimation of Boundaries between Speech Units Using Bayesian Changepoint Detectors ...................................................... 291
   R. Čmejla and P. Sovka

Data Driven Design of Filter Bank for Speech Recognition ............... 299
   L. Burget and H. Heřmanský

Minimization of Transition Noise and HNM Synthesis in Very Low Bit Rate Speech Coding .................................................. 305
   P. Motlíček, G. Baudoin, J. Černocký, and G. Chollet

Chinese Radicals and Top Ontology in EuroWordNet ........................ 313
   S.H.S. Wong and K. Pala
Towards a Dynamic Adjustment of the Language Weight .............................. 323  
  *G. Stemmer, V. Zeissler, E. Nöth, and H. Niemann*

Speech Recognition Based on Feature Extraction  
with Variable Rate Frequency Sampling .............................................. 329  
  *I. Potamitis, N. Fakotakis, and G. Kokkinakis*

The Phonectic SMS Reader ............................................................... 334  
  *J. Gros, F. Mihelič, N. Pavešić, M. Žganec, A. Mihelič, M. Knez, A. Merčun, and D. Škerl*

### Dialogue

Research Issues for the Next Generation Spoken Dialogue Systems Revisited .... 341  
  *E. Nöth, M. Boros, J. Fischer, F. Gallwitz, J. Haas, R. Huber, H. Niemann, G. Stemmer, and V. Warnke*

Determining User Interface Semantics Using Communicating Agents ............. 349  
  *L. Ton, L.J.M. Rothkrantz*

Agent-Based Adaptive Interaction and Dialogue Management Architecture  
for Speech Applications ..................................................................... 357  
  *M. Turunen and J. Hakulinen*

Dialogue for Web Search Utilizing Automatically Acquired  
Domain Knowledge ............................................................................. 365  
  *U. Kruschwitz*

A Development Tool for Spoken Dialogue Systems and Its Evaluation ........ 373  
  *S. Kogure and S. Nakagawa*

Using XML for Representing Domain Dependent Knowledge in Dialogos ........ 381  
  *G. Boella, R. Damiano, M. Danieli, and L. Lesmo*

Dialogue Generation of Program Source Codes ..................................... 388  
  *P. Nygrén and R. Batůšek*

Semi-automatic Identification of Danish Discourse Deictics ....................... 396  
  *C. Navarretta*

Language Understanding Using Two-Level Stochastic Models  
with POS and Semantic Units ......................................................... 403  
  *F. Pla, A. Molina, E. Sanchis, E. Segarra, F. García*

Shallow Processing and Cautious Incrementality in a Dialogue System Front End:  
Two Steps towards Robustness and Reactivity .................................. 410  
  *T. Lager*
Creation of a Corpus of Training Sentences
Based on Automated Dialogue Analysis ............................. 418
J. Schwarz and V. Matoušek

Dialogue Manager in Dialogue System for Visually Impaired Programmers .... 427
K. Taušer, V. Matoušek, and R. Mouček

Help System in Dialogue Grammar Based Programming Environment
for Visually Impaired Programmers .................................... 435
R. Mouček, K. Taušer, and P. Nygrýn

Automatic Generation of Dialogue Interfaces for Web-Based Applications ...... 443
L. Bártek

Author Index ................................................................. 451