

# Table of Contents

---

## I Invited Papers

---

Speech and Language Processing: Can We Use the Past to Predict the Future? . . . . .	3
<i>Kenneth Church (Microsoft, USA)</i>	
Common Sense About Word Meaning: Sense in Context . . . . .	15
<i>Patrick Hanks (Berlin-Brandenburg Academy of Sciences, Germany), James Pustejovsky (Brandeis University, USA)</i>	
ScanSoft's Technologies . . . . .	19
<i>Jan Odijk (ScanSoft Belgium)</i>	

---

## II Text

---

A Positional Linguistics-Based System for Word Alignment . . . . .	23
<i>Ana-Maria Barbu (Romanian Academy, Bucharest, Romania)</i>	
Handling Multi-word Expressions Without Explicit Linguistic Rules in an MT System . . . . .	31
<i>Akshar Bharati, Rajeev Sangal, Dipti Mishra, Sriram Venkatapathy, Papi Reddy T. (International Institute of Information Technology, Hyderabad, India)</i>	
The Szeged Corpus: A POS Tagged and Syntactically Annotated Hungarian Natural Language Corpus . . . . .	41
<i>Dóra Csendes, János Csirik, Tibor Gyimóthy (University of Szeged, Hungary)</i>	
Item Summarization in Personalisation of News Delivery Systems . . . . .	49
<i>Alberto Díaz, Pablo Gervás (Universidad Complutense de Madrid, Spain)</i>	
IR-n System, a Passage Retrieval Architecture . . . . .	57
<i>Fernando Llopis, Héctor García Puigcerver, Mariano Cano, Antonio Toral, Héctor Espí (University of Alicante, Spain)</i>	
Event Clustering in the News Domain . . . . .	65
<i>Cormac Flynn, John Dunnion (University College Dublin, Ireland)</i>	
HANDY: Sign Language Synthesis from Sublexical Elements Based on an XML Data Representation . . . . .	73
<i>László Havasi (PannonVision, Szeged, Hungary), Helga M. Szabó (National Association of the Deaf, Budapest, Hungary)</i>	
Using Linguistic Resources to Construct Conceptual Graph Representation of Texts . . . . .	81
<i>Svetlana Hensman, John Dunnion (University College Dublin, Ireland)</i>	

Slovak National Corpus . . . . .	89
<i>Alexander Horák, Lucia Gianitsová, Mária Šimková, Martin Šmotlák, Radovan Garabík (Slovak Academy of Sciences Bratislava, Slovakia)</i>	
Grammatical Heads Optimized for Parsing and Their Comparison with Linguistic Intuition . . . . .	95
<i>Vladimír Kadlec, Pavel Smrž (Masaryk University in Brno, Czech Republic)</i>	
How Dominant Is the Commonest Sense of a Word? . . . . .	103
<i>Adam Kilgarriff (Lexicography MasterClass Ltd. and ITRI, University of Brighton, UK)</i>	
POS Tagging of Hungarian with Combined Statistical and Rule-Based Methods . . . . .	113
<i>András Kuba, András Hóczá, János Csirik (University of Szeged, Hungary)</i>	
Grammatical Relations Identification of Korean Parsed Texts Using Support Vector Machines . . . . .	121
<i>Songwook Lee, Jungyun Seo (Sogang University, Seoul, Korea)</i>	
Clustering Abstracts Instead of Full Texts . . . . .	129
<i>Pavel Makagonov (Mixteca University of Technology, Mexico), Mikhail Alexandrov (National Polytechnic Institute, Mexico), Alexander Gelbukh (National Polytechnic Institute, Mexico)</i>	
Bayesian Reinforcement for a Probabilistic Neural Net Part-of-Speech Tagger . . . . .	137
<i>Manolis Maragoudakis, Todor Ganchev, Nikos Fakotakis (University of Patras, Greece)</i>	
Automatic Language Identification Using Phoneme and Automatically Derived Unit Strings . . . . .	147
<i>Pavel Matějka (FEEC VUT Brno, Czech Republic), Igor Szöke (FIT VUT Brno, Czech Republic and ESIEE Paris, France), Petr Schwarz (FIT VUT Brno, Czech Republic), and Jan Černocký (FIT VUT Brno, Czech Republic)</i>	
Slovak Text-to-Speech Synthesis in ARTIC System . . . . .	155
<i>Jindřich Matoušek, Daniel Tihelka (University of West Bohemia in Pilsen, Czech Republic)</i>	
Identifying Semantic Roles Using Maximum Entropy Models . . . . .	163
<i>Paloma Moreda, Manuel Fernández, Manuel Palomar, Armando Suárez (University of Alicante, Spain)</i>	
A Lexical Grammatical Implementation of Affect . . . . .	171
<i>Matthijs Mulder (University of Twente, Enschede, The Netherlands and Parabots Services, Amsterdam, The Netherlands) Anton Nijholt (University of Twente, Enschede, The Netherlands), Marten den Uyl, Peter Terpstra (Parabots Services, Amsterdam, The Netherlands)</i>	

Towards Full Lexical Recognition . . . . .	179
<i>Gordana Pavlović-Lažetić, Duško Vitas, Cvetana Krstev (University of Belgrade)</i>	
Discriminative Models of SCFG and STSG . . . . .	187
<i>Antoine Rozenknop, Jean-Cédric Chappelier, Martin Rajman (LIA, IIF, IC, EPFL, Lausanne, Switzerland)</i>	
Coupling Grammar and Knowledge Base: Range Concatenation Grammars and Description Logics . . . . .	195
<i>Benoît Sagot (Université Paris 7 and INRIA, France), Adil El Ghali (Université Paris 7, France)</i>	
Balancing Manual and Automatic Indexing for Retrieval of Paper Abstracts . . . . .	203
<i>Kwangcheol Shin (Chung-Ang University, Korea), Sang-Yong Han (Chung-Ang University, Korea), Alexander Gelbukh (National Polytechnic Institute, Mexico)</i>	
Unsupervised Learning of Rules for Morphological Disambiguation . . . . .	211
<i>Pavel Šmerk (Masaryk University in Brno, Czech Republic)</i>	
Ambiguous Supertagging Using a Feature Structure . . . . .	217
<i>François Toussenet (University Paris 7, France)</i>	
A Practical Word Sense Disambiguation System with High Performance for Korean . .	225
<i>Yeohoon Yoon (ETRI, Republic of Korea), Songwook Lee (Sogang University, Seoul, Republic of Korea), Joochan Sohn (ETRI, Republic of Korea)</i>	
Morphological Tagging of Russian Texts of the XIX <sup>th</sup> Century . . . . .	235
<i>Victor Zakharov, Sergei Volkov (St. Petersburg State University, Russia)</i>	

---

### III Speech

---

Large Vocabulary Continuous Speech Recognition for Estonian Using Morphemes and Classes . . . . .	245
<i>Tanel Alumäe (Tallinn Technical University, Estonia)</i>	
A New Classifier for Speaker Verification Based on the Fractional Brownian Motion Process . . . . .	253
<i>Ricardo Sant Ana, Rosângela Coelho (Instituto Militar de Engenharia, Rio de Janeiro, Brazil), Abraham Alcaim (Pontifícia Universidade Católica do Rio de Janeiro, Brazil)</i>	
A Universal Human Machine Speech Interaction Language for Robust Speech Recognition Applications . . . . .	261
<i>Ebru Arısoy, Levent M. Arslan (Boğaziçi University, Istanbul, Turkey)</i>	
Embedded ViaVoice . . . . .	269
<i>Tomáš Beran, Vladimír Bergl, Radek Hampl, Pavel Krbec, Jan Šedivý, Bořivoj Tydlitát, Josef Vopička (IBM Research Prague, Czech Republic)</i>	

New Speech Enhancement Approach for Formant Evolution Detection . . . . .	275
<i>Jesus Bobadilla (U.P.M. Madrid, Spain)</i>	
Measurement of Complementarity of Recognition Systems . . . . .	283
<i>Lukáš Burget (VUT Brno, Czech Republic)</i>	
Text-to-Speech for Slovak Language . . . . .	291
<i>Peter Čáky, Martin Klimo, Igor Mihálik, Radovan Mladšík (University of Žilina, Slovakia)</i>	
Speaker Verification Based on Wavelet Packets . . . . .	299
<i>Todor Ganchev, Mihalis Sifarakis, Nikos Fakotakis (University of Patras, Greece)</i>	
A Decoding Algorithm for Speech Input Statistical Translation . . . . .	307
<i>Ismael García-Varea (Univ. de Castilla-La Mancha, Albacete, Spain), Alberto Sanchis, Francisco Casacuberta (Univ. Politécnica de Valencia, Spain)</i>	
Aggregation Operators and Hypothesis Space Reductions in Speech Recognition . . . .	315
<i>Gábor Gosztolya, András Kocsor (University of Szeged, Hungary)</i>	
Combinations of TRAP Based Systems . . . . .	323
<i>František Grézl (Brno University of Technology, Czech Republic and IDIAP, Switzerland)</i>	
Automatic Recognition and Evaluation of Tracheoesophageal Speech . . . . .	331
<i>Tino Haderlein, Stefan Steidl, Elmar Nöth, Frank Rosanowski, Maria Schuster (University Erlangen-Nürnberg, Germany)</i>	
Using Neural Networks to Model Prosody in Czech TTS System Epos . . . . .	339
<i>Petr Horák (Academy of Sciences, Prague, Czech Republic), Jakub Adámek (Charles University, Prague, Czech Republic), Daniel Sobe (Dresden University of Technology, Federal Republic of Germany)</i>	
Auditory Scene Analysis via Application of ICA in a Time-Frequency Domain . . . . .	347
<i>Ladislava Janků (Czech Technical University in Prague, Czech Republic and Technical University Brno, Czech Republic)</i>	
Using the Lemmatization Technique for Phonetic Transcription in Text-to-Speech System . . . . .	355
<i>Jakub Kanis, Luděk Müller (University of West Bohemia in Pilsen, Czech Republic)</i>	
Automatic Categorization of Voicemail Transcripts Using Stochastic Language Models	363
<i>Konstantinos Koumpis (Vienna Telecommunications Research Center - ftw., Austria)</i>	
Low Latency Real-Time Vocal Tract Length Normalization . . . . .	371
<i>Andrej Ljolje, Vincent Goffin, Murat Saraclar (AT&amp;T Labs, Florham Park, USA)</i>	
Multimodal Phoneme Recognition of Meeting Data . . . . .	379
<i>Petr Motlíček, Jan Černocký (FIT VUT Brno, Czech Republic)</i>	

A New Multi-modal Database for Developing Speech Recognition Systems for an Assistive Technology Application . . . . .	385
<i>António Moura (Polytechnic Institute of Bragança, Portugal), Diamantino Freitas, Vitor Pera (University of Porto, Portugal)</i>	
Obtaining and Evaluating an Emotional Database for Prosody Modelling in Standard Basque . . . . .	393
<i>Eva Navas, Inmaculada Hernández, Amaia Castelruiz, Iker Luengo (University of the Basque Country, Bilbao, Spain)</i>	
Fully Automated Approach to Broadcast News Transcription in Czech Language . . . .	401
<i>Jan Nouza, Jindřich Žďánský, Petr David (Technical University of Liberec, Czech Republic)</i>	
A Computational Model of Intonation for Yorùbá Text-to-Speech Synthesis: Design and Analysis . . . . .	409
<i>Ọdétúnjí A. Ọdẹjọbí, Anthony J. Beaumont, Shun Ha Sylvia Wong (Aston University, UK)</i>	
Dynamic Unit Selection for Very Low Bit Rate Coding at 500 bits/sec . . . . .	417
<i>Marc Padellini, Francois Capman (Thales Communication, Colombes, France), Geneviève Baudoin (ESIEE, Noisy-Le-Grand, France)</i>	
On the Background Model Construction for Speaker Verification Using GMM . . . . .	425
<i>Aleš Padrta, Vlasta Radová (University of West Bohemia in Pilsen, Czech Republic)</i>	
A Speaker Clustering Algorithm for Fast Speaker Adaptation in Continuous Speech Recognition . . . . .	433
<i>Luis Javier Rodríguez, M. Inés Torres (Universidad del País Vasco, Bilbao, Spain)</i>	
Advanced Prosody Modelling . . . . .	441
<i>Jan Romportl, Jindřich Matoušek, Daniel Tihelka (University of West Bohemia in Pilsen, Czech Republic)</i>	
Voice Stress Analysis . . . . .	449
<i>Leon J.M. Rothkrantz, Pascal Wiggers, Jan-Willem A. van Wees, Robert J. van Vark (Delft University of Technology, The Netherlands)</i>	
Slovak Speech Database for Experiments and Application Building in Unit-Selection Speech Synthesis . . . . .	457
<i>Milan Rusko, Marian Trnka, Sachia Daržágín, Miloš Cerňák (Slovak Academy of Sciences, Bratislava, Slovakia)</i>	
Towards Lower Error Rates in Phoneme Recognition . . . . .	465
<i>Petr Schwarz, Pavel Matějka, Jan Černocký (VUT Brno, Czech Republic)</i>	
Examination of Pronunciation Variation from Hand-Labelled Corpora . . . . .	473
<i>György Szaszák, Klára Vicsi (Budapest University for Technology and Economics, Hungary)</i>	

New Refinement Schemes for Voice Conversion . . . . .	481
<i>Abdelgawad Eb. Taher (Brno University of Technology, Czech Republic)</i>	
Acoustic and Linguistic Information Based Chinese Prosodic Boundary Labelling . . . .	489
<i>Jianhua Tao (Chinese Academy of Sciences, Beijing, China)</i>	
F0 Prediction Model of Speech Synthesis Based on Template and Statistical Method . .	497
<i>Jianhua Tao (Chinese Academy of Sciences, Beijing, China)</i>	
An Architecture for Spoken Document Retrieval . . . . .	505
<i>Rafael M. Terol, Patricio Martínez-Barco, Manuel Palomar (Universidad de Alicante, Spain)</i>	
Evaluation of the Slovenian HMM-Based Speech Synthesis System . . . . .	513
<i>Boštjan Vesnicher, France Mihelič (University of Ljubljana, Slovenia)</i>	
Modeling Prosodic Structures in Linguistically Enriched Environments . . . . .	521
<i>Gerasimos Xydias, Dimitris Spiliotopoulos, Georgios Kouroupetroglou (University of Athens, Greece)</i>	
Parallel Root-Finding Method for LPC Analysis of Speech . . . . .	529
<i>Juan-Luis García Zapata, Juan Carlos Díaz Martín (Universidad de Extremadura, Spain), Pedro Gómez Vilda (Universidad Politécnica de Madrid, Spain)</i>	
Automatic General Letter-to-Sound Rules Generation for German Text-to-Speech System . . . . .	537
<i>Jan Zelinka, Luděk Müller (University of West Bohemia in Pilsen, Czech Republic)</i>	
Pitch Accent Prediction from ToBI Annotated Corpora Based on Bayesian Learning . .	545
<i>Panagiotis Zervas, Nikos Fakotakis, George Kokkinakis (University of Patras, Greece)</i>	
Processing of Logical Expressions for Visually Impaired Users . . . . .	553
<i>Pavel Žikovský (Czech Technical University in Prague, Czech Republic), Tom Pešina (Charles University in Prague, Czech Republic), Pavel Slavík (Czech Technical University in Prague, Czech Republic)</i>	

---

## IV Dialogue

---

Durational Aspects of Turn-Taking in Spontaneous Face-to-Face and Telephone Dialogues . . . . .	563
<i>Louis ten Bosch, Nelleke Oostdijk (Nijmegen University, The Netherlands), Jan Peter de Ruiter (Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands)</i>	
A Speech Platform for a Bilingual City Information System . . . . .	571
<i>Thomas Brey (University of Regensburg, Germany), Tomáš Pavelka (University of West Bohemia in Pilsen, Czech Republic)</i>	

Rapid Dialogue Prototyping Methodology . . . . .	579
<i>Trung H. Bui, Martin Rajman, Miroslav Melichar (EPFL, Lausanne, Switzerland)</i>	
Building Voice Applications from Web Content . . . . .	587
<i>César González-Ferreras, Valentín Cardeñoso-Payo (Universidad de Valladolid, Spain)</i>	
Information-Providing Dialogue Management . . . . .	595
<i>Melita Hajdinjak, France Mihelič (University of Ljubljana, Slovenia)</i>	
Realistic Face Animation for a Czech Talking Head . . . . .	603
<i>Zdeněk Krňoul, Miloš Železný (University of West Bohemia in Pilsen, Czech Republic)</i>	
Evaluation of a Web Based Information System for Blind and Visually Impaired Students: A Descriptive Study . . . . .	611
<i>Stefan Riedel, Wolfgang Wünschmann (Dresden University of Technology, Germany)</i>	
Multimodal Dialogue Management . . . . .	621
<i>Leon J.M. Rothkrantz, Pascal Wiggers, Frans Flippo, Dimitri Woei-A-Jin, Robert J. van Vark (Delft University of Technology, The Netherlands)</i>	
Looking at the Last Two Turns, I'd Say This Dialogue Is Doomed – Measuring Dialogue Success . . . . .	629
<i>Stefan Steidl, Christian Hacker, Christine Ruff, Anton Batliner, Elmar Nöth (University Erlangen-Nürnberg, Germany), Jürgen Haas (Sympalog Voice Solutions GmbH, Erlangen, Germany)</i>	
Logical Approach to Natural Language Understanding in a Spoken Dialogue System .	637
<i>Jeanne Villaneau (Université de Bretagne-Sud), Jean-Yves Antoine (Université de Bretagne-Sud), Olivier Ridoux (Université de Rennes 1)</i>	
Building a Dependency-Based Grammar for Parsing Informal Mathematical Discourse	645
<i>Magdalena Wolska, Ivana Kruijff-Korbayová (Saarland University, Saarbrücken, Germany)</i>	
<b>Colophon</b> . . . . .	653
<b>Subject Index</b> . . . . .	655
<b>Author Index</b> . . . . .	665