Table of Contents

I Text

TreeTalk-D: A Machine Learning Approach to Dutch Word Pronunciation 3
Bertjan Busser (Tilburg University)

Czech National Corpus: Its Character, Goal and Background 9
František Čermák (Charles University, Prague)

Lexithiras: Multi-lingual Corpus Lexicography on PCs 15
Evangelos Dermatas, George Kokkinakis (University of Patras)

Text Understanding and Interpreting Methodologies Based on Hermeneutic and Linguistic Techniques 21
Sergey V. Chebanov (Russian Academy of Science), Gregory Y. Martynenko, Tatiana Y. Sherstinova (St. Petersburg State University)

EULER: Multi-Lingual Text-to-Speech Project 27
T. Dutoit, F. Malfrère, V. Pagel (Faculte Polytechnique de Mons) P. Mertens (Kath. University Leuven) M. Bagein, A. Ruelle, A. Gilman (Faculte Polytechnique de Mons)

Developing a Language Independent Interactive Bible Concordance on the World Wide Web 33
J. Gareth Evans (Swansea Institute of Higher Education)

Tagging a Morphologically Rich Language 39
Marcelo Finger (University of Sao Paulo)

Prague Dependency Treebank: From Analytic to Tectogrammatical Annotations . 45
Eva Hajičová (Charles University, Prague)

Epos — A New Approach to the Speech Synthesis 51
Jiří Hanika (Charles University, Prague), Petr Horák (Academy of Sciences, Prague)

Technical Insights into the Birth of the Czech National Corpus 55
Jaroslava Hlaváčová (Charles University, Prague)

Verb Valency and Semantic Classification of Verbs 61
Aleš Horák (Masaryk University, Brno)

Automatic Generation of Instructions in a Multiiingual Environment 67
Ivana Kruijff-Korbayová (Charles University, Prague)

Adaptation of a Lexical Database in Text Categorization Task 73
L. Alfonso Ureña (Universidad de Jaén), M. De Buenaga (Universidad Europea de Madrid), M. García (Universidad de Jaén), J. M. Gómez (Universidad Europea de Madrid)

Connection Driven Parsing of Lexicalized TAG 79
Patrice Lopez (University of Nancy)

Construction of Structural Thematic Summary of Text 85
Natalia V. Loukachevitch, Boris V. Dobrov (Moscow State University)

Using Local Grammars for Agreement Modeling in Highly Inflective Languages . 91
Goran Nenadić, Duško Vitas (University of Belgrade)

Mutual Information in Czech Corpus ESO 97
Karel Pala, Pavel Rychlý (Masaryk University, Brno)

Recent Improvements in Slovene Text-to-Speech System 103
Tomaz Sef, Ales Dobnikar, Matjaz Gams, Zeljko Khermayer (Jozef Stefan Institute, Ljubljana)

The Improvement of Common Statistical Measure 109
Pavel Rychlý (Masaryk University, Brno)

Partial Parsing Method Applied to Rules Acquisition for Medical Expert System . 113
Maciej Piasecki, Jerzy Sas (Wrocław University of Technology)

TENCO — Automatic Text Encoder 119
Milena Slavcheva, Boyanka Zaharieva (Bulgarian Academy of Sciences; Sofia)

Semiautomatic Verb Valencies Acquisition and Checking 123
Pavel Smrž, Aleš Horák (Masaryk University, Brno)

New Tools for Disambiguation of Czech Texts 129
Pavel Smrž, Eva Žáčková (Masaryk University, Brno)

Optimal Joint Procedure for Current Pitch Period Discrimination and Speech Signal Partition into Quasi-Periodic and Non-Periodic Segments 135
Taras K. Vintsiuk (NAS Institute of Cybernetics & UNESCO/IIP International Research-Training Centre for Information Technologies and Systems, Kyjiv)

How Much Training Data Is Required to Remove Data Sparseness in Statistical Language Learning? 141
Dan-Hee Yang, Mansuk Song (Yonsei University, Seoul)

Classification of Medical Unstructured Text Documents Using the Naïve Bayes Algorithm 147
Jan Žižka, Aleš Bourek (Masaryk University, Brno)

II Speech

Statistics of the Syllable Segments for Speech Synthesis of the Czech Language . 153
Robert Batůšek (Masaryk University, Brno)

Confidence Measures in Hybrid HMM/ANN Speech Recognition 159
Giulia Bernardis (Dalle Molle Institute for Perceptual Artificial Intelligence, Martigny), Hervé Bourlard (Swiss Federal Institute of Technology, Lausanne)

Experience from the Field — Name Dialing en Masse 165
Vladimír Bergl, Ken Davies, Abraham Ittycheriah, Andrzej Sakrajda

Harmonics Tracking and Sieving-Integrating Algorithm for Robust Speech Recognition 171
Zhang Bo (City University of Hongkong)

A Neural Network-Based Speaker-Independent System for Word Recognition in Romanian Language 177
Dragoş Burileanu, Mihai Sima, Corneliu Burileanu, Victor Croitoru (University of Bucharest)

Automatically Derived Speech Units: Applications to Very Low Rate Coding and Speaker Verification 183
Jan Černocký (VUT Brno), Geneviève Baudoin (ESIEE Paris), Dijana Petrovska-Delacrétaz (DE–CIRC, EPFL Lausanne), Jean Hennebert (DE–CIRC, EPFL Lausanne), Gérard Chollet (ENST Paris)

Cepstral Trajectory Transformation for Subword Recognition 189
Tibor Fegyó, Péter Tatai (Technical University of Budapest)

The Romanian ROMVOX Text-to-Speech Synthesis System Using a Mixed Time Domain-LPC Pitch Alteration Technique 195
Attila Ferencz, István Nagy, Tünde-Csilla Kovács, Maria Ferencz, Teodora Raţiu (Software ITC), Diana Zaiu (Technical University of Cluj-Napoca)

Pushing Forward the Interface between Recognition and Understanding — How to Integrate Syntactic Structure into the Output of a Word Recognizer 201
F. Gallwitz, A. Batliner, J. Buckow, R. Huber, H. Niemann, E. Nöth (Universität Erlangen-Nürnberg)

Interfacing of CASA and Multistream Recognition 207
Hervé Glotin (ICP Grenoble, IDIAP Martigny), Frédéric Berthommier, Emmanuel Tessier (ICP Grenoble), Hervé Bourlard (IDIAP Martigny)

Data-Driven Speech Analysis For ASR 213
Hynek Hermansky (Oregon Graduate Institute of Science & Technology, ICSI, Berkeley, Technical University, Brno)

NÁRODNÍ KNIHOVNA ČESKÉ REPUBLIKY UKF

The LPC Analysis and Synthesis of F0 Contour 219
Petr Horák (Academy of Sciences, Prague)

You BEEP Machine — Emotion in Automatic Speech Understanding Systems ... 223
R. Huber, E. Nöth, A. Batliner, J. Buckow, V. Warnke, H. Niemann (University of Erlangen-Nuremberg)

An Overview of the Slovenian Spoken Dialog System 229
Ivo Ipšić, France Mihelič, Simon Dobrišek, Jerneja Gros, Nikola Pavešić (Ljubljana University)

Detection of Sentence Types by the Integrated Prosody Module 235
Jana Klečková, Václav Matoušek (University of West Bohemia, Plzeň)

Extraction of Intonation in Thai Language by Karhunen-Loeve Transformation .. 241
Rachada Kongkachandra, Kreingsak Tamee, Chom Kimpan (King's Mongkut Institute of Technology, Bangkok)

Speech Synthesis Based on the Composed Syllable Segments 247
Ivan Kopeček (Masaryk University, Brno)

Speech Recognition in a Real Room and Multi-Simultaneous-Speaker Environment 251
Athanasios Koutras, Evangelos Dermatas, George Kokkinakis (University of Patras)

Design and Implementation of a Dialog Manager 257
Jana Krutišová, Václav Matoušek, Jana Ocelíková (University of West Bohemia, Plzeň)

Phonetic Transcription for the Systems of Speech Synthesis and Recognition 263
Vladimir I. Kuznetsov, Tatiana Y. Sherstinova (St. Petersburg University)

Generating Spoken Utterances from Concepts and Prosodic Schemes 269
Pierre Larrey, Nadine Vigouroux, Guy Pérennou (IRIT)

Speaker Verification Using Vector Quantisation 275
Eugen Lupu, Petre Pop, Gavril Toderean (Technical University of Cluj-Napoca)

Speech Recognition in Noisy Reverberant Rooms Using Frequency Domain Adaptive Filtering 281
George Nokas, Evangelos Dermatas, George Kokkinakis (University of Patras)

Automatic Phonemic Segmentation of Brazilian Portuguese Speech Databases .. 287
Henrique F. Nunes (Zetax Tecnologia Brasil), Edson J. Nagle, Cairo H. da Silva, Fernando Runstein, (Fundação Centro de Pesquisa e Desenvolvimento en Telecomunicações, Brasil)

Recognition of a Weather Forecast Transmitted by the Czech Radio Broadcasting 293
Luděk Müller, Josef Psutka (University of West Bohemia in Plzeň)

Design of the Czech National Speech Corpus for Speech Recognition Applications with a Large Vocabulary 299
Vlasta Radová (University of West Bohemia in Plzeň)

Dutch Automatic Speech Recognition 305
J. Schalken, L. J. M. Rothkrantz (Delft University of Technology)

Development of Multilingual Acoustic Models in the *GlobalPhone* Project 311
Tanja Schultz, Alex Waibel (Universität Karlsruhe, Carnegie Mellon University)

Using Rules of Consonant Distribution for Russian Continuous Speech Automatic Segmentation 317
P. Skrelin, K. Shalonova (St. Petersburg State University)

Validity Criterion for Unsupervised Speaker Recognition 321
Itshak Voitovetsky, Hugo Guterman, Arnon Cohen (Ben-Gurion University of the Negev)

The Structure of the Multimedia Database STDB 327
Ralf Vollmann (University of Graz)

Diphone Voice Synthesis Using Fast Fourier Transformation 333
Pavel Žikovský and Pavel Slavík (Czech Technical University, Prague)

Simple Speech Classification Using Automatically Generated Rules 339
Jan Žižka (Masaryk University, Brno), Zdeněk Kratochvíl (Technical University, Brno), Luboš Popelínský (Masaryk University, Brno)

III Dialogue

Using Dialog Programming for Developing the Speech Oriented Hypertext System AUDIS 347
Luděk Bártek, Robert Batůšek, Marek Fikera, Ivan Kopeček (Masaryk University, Brno)

An Approach of Relevance Theory Modeling in Question Answering Systems ... 351
Guylaine Gonel (Universite Paris 13), Bernard Levrat (LERIA, Paris)

A Concept for a Prosodically and Statistically Driven Chunky Semantic Parser .. 357
Jürgen Haas, Manuela Boros, Elmar Nöth, Volker Warnke, Heinrich Niemann (University of Erlangen-Nürnberg)

Using TIL for Semantic Analysis of Text 363
Leo Hadacz (Masaryk University, Brno)

Partial Parsing and Repairs for Spoken Input 369
Ariane Halber (ENST Paris, Thomson-CSF)

Multilingual Speech Recognition in the Context of Multilingual Information Retrieval Dialogues 375
Stefan Harbeck, Elmar Nöth, Heinrich Niemann (University of Erlangen-Nürnberg)

Textual Deep Structure 381
Viatcheslav Iatsko (State University of Khakasia)

Developing a Model of Dialog Strategy 387
Mare Koit, Haldur Õim (University of Tartu)

Dialog Based Programming 391
Ivan Kopeček (Masaryk University, Brno)

"Natural Speaking" Information Retrieval Dialog System 397
Jana Krutišová, Jana Oceliková, Václav Matoušek, Pavel Mautner (University of West Bohemia, Plzeň)

Users' Behaviours in Spontaneous Oral Dialogue Strategy Design 403
Carine-Alexia Lavelle, Martine de Calmès, Guy Pérennou (IRIT)

A Voice Dialing System for Mobile Phones 409
Bálint Lükö (Technical University of Budapest)

A Natural Language Web-based Dialogue System with a Talking Face 415
Anton Nijholt, Mathieu van den Berk, Arjan van Hessen (University of Twente)

Generating Intonation Contours Using Tonal Specifications 421
Hannes Pirker, Erhard Rank (Öfai Vienna), Harald Trost (University of Vienna)

An Interface Language for Dialogues 427
Violeta Quental, Maria Carmelita Padua Dias (Pontifícia Universidade Católica do Rio de Janeiro), Laura Sanchez Garcia (Universidade Federal do Paraná)

An Automated Speech Recognition System for Fast-food Restaurant Applications 433
L. J. M. Rothkrantz, F. Mohamed-Hoesein, A. Shirzad (Delft University of Technology)

Remarks on Parsing Written and Spoken Discourse 439
Petr Sgall (Charles University, Prague)

Author Index 445

Colophon 447