

Table of Contents

I Text

A Common Solution for Tokenization and Part-of-Speech Tagging	3
<i>Jorge Graña, Miguel A. Alonso, Manuel Vilares (University La Coruña)</i>	
Rule Parser for Arabic Stemmer	11
<i>Imad A. Al-Sughaiyer, Ibrahim A. Al-Kharashi (Computer and Electronics Research Institute, Saudi Arabia)</i>	
Achieving an Almost Correct PoS-Tagged Corpus	19
<i>Pavel Květoň (Charles University, Prague), Karel Oliva (Austrian Research Institute for Artificial Intelligence, Vienna)</i>	
Evaluation of a Japanese Sentence Compression Method Based on Phrase Significance and Inter-Phrase Dependency	27
<i>Rei Oguro, Hiromi Sekiya, Yuhei Morooka, Kazuyuki Takagi, Kazuhiko Ozeki (The University of Electro-Communications, Tokyo)</i>	
User Query Understanding by the InBASE System as a Source for a Multilingual NL Generation Module	33
<i>Michael V. Boldasov (Moscow State University), Elena G. Sokolova (Russian Research Institute for Artificial Intelligence), Michael G. Malkovsky (Moscow State University)</i>	
The Role of WSD for Multilingual Natural Language Applications	41
<i>Andrés Montoyo, Rafael Romero, Sonia Vázquez, Carmen Calle, Susana Soler (University Alicante)</i>	
Gibbsian Context-Free Grammar for Parsing	49
<i>Antoine Rozenknop (Swiss Federal Institute of Technology, Lausanne)</i>	
Cross-Language Access to Recorded Speech in the MALACH Project	57
<i>D. W. Oard, D. Demner-Fushman (University of Maryland), J. Hajič (Charles University, Prague), B. Ramabhadran (IBM T. J. Watson Research Center), S. Gustman (Survivors of the Shoah Visual History Foundation, Los Angeles), W. J. Byrne (Johns Hopkins University), D. Soergel, B. Dorr, P. Resnik (University of Maryland) and M. Picheny (IBM T. J. Watson Research Center)</i>	
Using Salient Words to Perform Categorization of Web Sites	65
<i>Marek Trabalka, Mária Bieliková (Slovak University of Technology, Bratislava)</i>	
Discourse-Semantic Analysis of Hungarian Sign Language	73
<i>Gábor Alberti (University of Pécs), Helga M. Szabó (National Association of the Deaf, Hungary)</i>	

VIII Table of Contents

Dependency Analyser Configurable by Measures	81
<i>Tomáš Holan (Charles University, Prague)</i>	
The Generation and Use of Layer Information in Multilayered Extended Semantic Networks	89
<i>Sven Hartrumpf, Hermann Helbig (University of Hagen)</i>	
Filtering of Large Numbers of Unstructured Text Documents by the Developed Tool TEA	99
<i>Jan Žižka, Aleš Bourek (Masaryk University in Brno)</i>	
Visualisation Techniques for Analysing Meaning	107
<i>Dominic Widdows, Scott Cederberg, Beate Dorow (Stanford University)</i>	
Statistical Part-of-Speech Tagging for Classical Chinese	115
<i>Liang Huang, Yinan Peng (Shanghai Jiaotong University), Huan Wang (East China Normal University), Zhenyu Wu (Shanghai Jiaotong University)</i>	
Spanish Natural Language Interface for a Relational Database Querying System	123
<i>Rodolfo A. Pazos R. (Centro Nacional de Investigación y Desarrollo Tecnológico, Mexico), Alexander Gelbukh (National Polytechnic Institute, Mexico), J. Javier González B., Erika Alarcón R., Alejandro Mendoza M., A. Patricia Domínguez S. (Instituto Tecnológico de Ciudad Madero, Mexico)</i>	
Word Sense vs. Word Domain Disambiguation: A Maximum Entropy Approach	131
<i>Armando Suárez, Manuel Palomar (University Alicante, Spain)</i>	
Exploiting Thesauri and Hierarchical Categories in Cross-Language Information Retrieval	139
<i>Fatiha Sadat, Masatoshi Yoshikawa, Shunsuke Uemura (Nara Institute of Science and Technology)</i>	
Valency Lexicon for Czech: From Verbs to Nouns	147
<i>Markéta Lopatková, Veronika Řezníčková, Zdeněk Žabokrtský (Charles University, Prague)</i>	
Term Clustering Using a Corpus-Based Similarity Measure	151
<i>Goran Nenadić, Irena Spasić, Sophia Ananiadou (University of Salford, UK)</i>	
Word Sense Discrimination for Czech	155
<i>Robert Král (Masaryk University in Brno)</i>	
Tools for Semi-Automatic Assignment of Czech Nouns to Declination Patterns	159
<i>Dita Bartůšková, Radek Sedláček (Masaryk University in Brno)</i>	

II Speech

Automatic Lexical Stress Assignment of Unknown Words for Highly Inflected Slovenian Language	165
<i>Tomaž Šef, Maja Škrjanc, Matjaž Gams (Ljubljana University)</i>	
German and Czech Speech Synthesis Using HMM-Based Speech Segment Database . .	173
<i>Jindřich Matoušek, Daniel Tihelka, Josef Psutka, Jana Hesová (University of West Bohemia in Pilsen)</i>	
Comparison and Combination of Confidence Measures	181
<i>Georg Stemmer, Stefan Steidl, Elmar Nöth, Heinrich Niemann, Anton Batliner (University Erlangen-Nürnberg)</i>	
Strategies for Developing a Real-Time Continuous Speech Recognition System for Czech Language	189
<i>Jan Nouza (Technical University of Liberec)</i>	
Comparative Study on Bigram Language Models for Spoken Czech Recognition	197
<i>Dana Nejedlová (Technical University of Liberec)</i>	
Integration of Speech Recognition and Automatic Lip-Reading	205
<i>Pascal Wiggers, Leon J. M. Rothkrantz (Delft University of Technology)</i>	
Utterance Verification Based on the Likelihood Distance to Alternative Paths	213
<i>Gies Bouwman, Lou Boves (University Nijmegen)</i>	
Rejection Technique Based on the Mumble Model	221
<i>Tomáš Bartoš, Luděk Müller (University of West Bohemia in Pilsen)</i>	
Efficient Noise Estimation and Its Application for Robust Speech Recognition	229
<i>Petr Motlíček, Lukáš Burget (Oregon Health & Science University and Brno University of Technology)</i>	
AlfaNum System for Speech Synthesis in Serbian Language	237
<i>Milan Sečujski, Radovan Obradović, Darko Pekar, Ljubomir Jovanov, Vlado Delić (University of Novi Sad)</i>	
Speech Features Extraction Using Cone-Shaped Kernel Distribution	245
<i>Janez Žibert, France Mihelič, Nikola Pavešić (University of Ljubljana)</i>	
Automatic Transcription of Czech Language Oral History in the MALACH Project: Resources and Initial Experiments	253
<i>Josef Psutka, Pavel Ircing, Josef V. Psutka, Vlasta Radová (University of West Bohemia in Pilsen), William J. Byrne (Johns Hopkins University), Jan Hajič (Charles University, Prague), Samuel Gustman (Survivors of the Shoah Visual History Foundation), Bhuvana Ramabhadran (IBM T.J. Watson Research Laboratory)</i>	

On the First Greek-TTS Based on Festival Speech Synthesis	261
<i>Zervas P., Potamitis I., Fakotakis N., Kokkinakis G. (University of Patras)</i>	
An Analysis of Limited Domains for Speech Synthesis	265
<i>Robert Batůšek (Masaryk University in Brno)</i>	
Advances in Very Low Bit Rate Speech Coding Using Recognition and Synthesis Techniques	269
<i>Geneviève Baudoin (ESIEE Paris), François Capman (Thales Communications), Jan Černocký (VUT Brno), Fadi El Chami (ESIEE Paris, University Tripolis), Maurice Charbit, Gérard Chollet (ENST Paris), Dijana Petrovska-Delacrétaz (University of Fribourg)</i>	
A Comparison of Different Approaches to Automatic Speech Segmentation	277
<i>Kris Demuyne, Tom Laureys (University Leuven)</i>	
Keyword Spotting Using Support Vector Machines	285
<i>Yassine Ben Ayed, Dominique Fohr, Jean Paul Haton (LORIA-CNRS/ INRIA Lorraine), Gérard Chollet (ENST, CNRS-LTCI)</i>	
Automatic Parameter Estimation for a Context- Independent Speech Segmentation Algorithm	293
<i>Guido Aversano (Universita Salerno and IIASS), Anna Esposito (Wright State University and IIASS)</i>	
Phoneme Lattice Based A* Search Algorithm for Speech Recognition	301
<i>Pascal Nocera, Georges Linares, Dominique Massonié, Loïc Lefort (LIA Avignon)</i>	
Heuristic and Statistical Methods for Speech/Non-Speech Detector Design	309
<i>Michal Prcín, Luděk Müller (University of West Bohemia)</i>	
An Analysis of Conditional Responses in Dialogue	317
<i>Elena Karajosova, Ivana Kruijff-Korbayová (Saarland University)</i>	
Some Like It Gaussian	321
<i>P. Matějka (FEEC VUT Brno), P. Schwarz (FIT VUT Brno), M. Karafiát (FIT VUT Brno) and J. Černocký (FIT VUT Brno)</i>	
Kernel Springy Discriminant Analysis and Its Application to a Phonological Awareness Teaching System	325
<i>András Kocsor (Hungarian Academy of Sciences), Kornél Kovács (University of Szeged)</i>	
Large Vocabulary Speech Recognition of Slovenian Language Using Data-Driven Morphological Models	329
<i>Tomaž Rotovnik, Mirjam Sepesy Maučec, Bogomir Horvat, Zdravko Kačič (University of Maribor)</i>	

Uniform Speech Recognition Platform for Evaluation of New Algorithms	333
<i>Andrej Žgank, Tomaž Rotovnik, Zdravko Kačič, Bogomir Horvat (University of Maribor)</i>	
Speech Enhancement Using Mixtures of Gaussians for Speech and Noise	337
<i>Ilyas Potamitis, Nikos Fakotakis, Nikos Liolios, George Kokkinakis (University of Patras)</i>	
Fitting German into N-Gram Language Models	341
<i>Robert Hecht, Jürgen Riedler, Gerhard Backfried (SAIL LABS Technology AG, Vienna)</i>	
Audio Collections of Endangered Arctic Languages in the Russian Federation	347
<i>Marina Lublinskaya, Tatiana Sherstinova (St. Petersburg State University)</i>	

III Dialogue

Prosodic Classification of Offtalk: First Experiments	357
<i>Anton Batliner, Viktor Zeißler, Elmar Nöth, Heinrich Niemann (University Erlangen-Nürnberg)</i>	
Statistical Decision Making from Text and Dialogue Corpora for Effective Plan Recognition	365
<i>Manolis Maragoudakis, Aristomenis Thanopoulos, Nikos Fakotakis (University of Patras)</i>	
Natural Language Guided Dialogues for Accessing the Web	373
<i>Marta Gattus, Horacio Rodríguez (Technical University of Catalonia)</i>	
Evaluating a Probabilistic Dialogue Model for a Railway Information Task	381
<i>Carlos D. Martínez-Hinarejos, Francisco Casacuberta (Valencia University)</i>	
Applying Dialogue Constraints to the Understanding Process in a Dialogue System . . .	389
<i>Emilio Sanchis, Fernando García, Isabel Galiano, Encarna Segarra (Universitat de València)</i>	
Evaluation of Prediction Methods Applied to an Inflected Language	397
<i>Nestor Garay-Vitoria, Julio Abascal, Luis Gardezabal (University of the Basque Country)</i>	
Knowledge Based Speech Interfacing for Handhelds	405
<i>C.K. Yang, L.J.M. Rothkrantz (Delft University of Technology)</i>	
Different Approaches to Build Multilingual Conversational Systems	413
<i>Marion Mast, Thomas Roß, Henrik Schulz (IBM Europe), Heli Harrikari (NOKIA Research Center)</i>	

XII Table of Contents

Strategies to Overcome Problematic Input in a Spanish Dialogue System	421
<i>Victoria Arranz, Núria Castell, Montserrat Cívít (Universitat Politècnica de Catalunya)</i>	
Dialogue Systems and Planning	429
<i>Guy Camilleri (University Toulouse)</i>	
A Flexible Framework for Evaluation of New Algorithms for Dialogue Systems	437
<i>Pavel Cenek (Masaryk University in Brno)</i>	
From HTML to VoiceXML: A First Approach	441
<i>César González Ferreras, David Escudero Mancebo, Valentín Cardeñoso Payo (University Valladolid)</i>	
Voice Chat with a Virtual Character: The Good Soldier Švejk Case Project	445
<i>Jan Nouza, Petr Kolář, Josef Chaloupka (Technical University of Liberec)</i>	
Application of Spoken Dialogue Technology in a Medical Domain	449
<i>I. Azzini, T. Giorgino (University of Pavia), D. Falavigna, R. Gretter (ITC-irst, Trento)</i>	
A Voice-Driven Web Browser for Blind People	453
<i>Simon Dobrišek, Jerneja Gros, Boštjan Vesnicer, France Mihelič, Nikola Pavešić (University of Ljubljana)</i>	
Enhancing Best Analysis Selection and Parser Comparison	461
<i>Aleš Horák, Vladimír Kadlec, Pavel Smrž (Masaryk University in Brno)</i>	
Author Index	467
Subject Index	471