Proceedings of the Paralinguistic Information and its Integration in Spoken Dialogue Systems Workshop
Proceedings of the Paralinguistic Information and its Integration in Spoken Dialogue Systems Workshop
The 3rd International Workshop on Spoken Dialogue Systems (IWSDS2011) was held at Granada, Spain, 1-3 September 2011, as a satellite event of Interspeech 2011. This annual workshop brings together researchers from all over the world working in the field of spoken dialogue systems. It provides an international forum for the presentation of research and applications, and for lively discussions among researchers as well as industrialists.

Following the success of IWSDS2009 (in Irsee, Germany) and IWSDS2010 (in Gotemba Kogen Resort, Japan), IWSDS2011 has designated "Paralinguistic Information and its Integration in Spoken Dialogue Systems" as a special theme for discussion, considering three Special Tracks:

- Spoken dialogue systems for robotics.
- Emotions and spoken dialogue systems.
- Spoken dialogue systems for real-world applications.

We also encourage discussions on common issues of spoken dialogue systems, including but not limited to:

- Speech recognition and semantic analysis.
- Dialogue management.
- Recognition of emotions from speech, gestures, facial expressions and physiological data.
- User modelling.
- Planning and reasoning capabilities for coordination and conflict description.
- Conflict resolution in complex multi-level decisions.
- Multi-modality such as graphics, gesture and speech for input and output.
- Fusion and information management.
- Learning and adaptability.
- Visual processing and recognition for advanced human-computer interaction.
- Databases and corpora.
- Evaluation strategies and paradigms.
- Prototypes and products.
The workshop programme consists of 35 papers and two invited keynote talks by Prof. Kristiina Jokinen and Prof. Roger K. Moore.

We thank the Scientific Committee members for their efficient contributions and for completing the review process on time.

Moreover, we express our gratitude to the Steering Committee for their guidance and suggestions, and to the Local Committee for their help with the organisational arrangements. In addition, we thank the Organising Committee of IWSD2009 for their support in updating the workshop’s web page, and the Organising Committee of IWSDS2010 for sharing with us their knowledge on local arrangements.

Furthermore, we must mention that this workshop would not have been possible without the support of:

- ACL (Association for Computational Linguistics).
- Dept. of Languages and Computer Systems, University of Granada, Spain.
- ISCA (International Speech Communication Association).
- Korean Society of Speech Scientists.
- SIGDIAL (ACL Special Interest Group on Discourse and Dialogue).
- University of Granada, Spain.

Last, but not least, we thank the authors for their contributions.

We hope all the attendees benefit from the workshop and enjoy their stay in Granada.

Granada, Spain
Tokyo, Japan

Ramón López-Cózar
Tetsunori Kobayashi
September 2011
IWSDS2011 was organised by the Dept. of Languages and Computer Systems (University of Granada, Spain) and the Dept. of Computer Science and Engineering (Waseda University, Japan).

**Chairs**

Ramón López-Cózar  
Tetsunori Kobayashi

**Local Committee**

Zoraida Callejas  
David Griol  
Gonzalo Espejo  
Nieves Ábalos  
Research Group on Spoken and Multimodal Dialogue Systems (SISDIAL)

**Scientific Committee**

Jan Alexandersson - DFKI, Saarbrücken, Germany  
Masahiro Araki - Interactive Intelligence lab, Kyoto Institute of Technology, Japan  
André Berton - Daimler R&D, Ulm, Germany  
Heriberto Cuayáhuítl - DFKI, Saarbrücken, Germany  
Sadaoki Furui - Tokyo Institute of Technology, Tokyo, Japan  
Joakim Gustafson - KTH, Stockholm, Sweden  
Tobias Heinroth - Ulm University, Germany  
Paul Heisterkamp - Daimler Research, Ulm, Germany  
Kristiina Jokinen - University of Helsinki, Finland  
Tatsuya Kawahara - Kyoto University, Japan  
Hong Kook Kim - Gwangju Institute of Science and Technology, Korea  
Lin-Shan Lee - National Taiwan University, Taiwan
Mike McTear - University of Ulster, UK
Mikio Nakano - Honda Research Institute, Japan
Elmar Nöth - University of Erlangen, Germany
Norbert Reithinger - DFKI, Berlin, Germany
Gabriel Skantze - KTH, Stockholm, Sweden
Alexander Schmitt - Ulm University, Germany
Kazuya Takeda - Graduate School of Information Science, Nagoya University, Japan
Hsin-min Wang - Academia Sinica, Taiwan

Steering Committee

Gary Geunbae Lee - POSTECH, Pohang, Korea
Joseph Mariani - LIMSI-CNRS and IMMI, Orsay, France
Wolfgang Minker - Ulm University, Germany
Satoshi Nakamura - NICT, Kyoto, Japan
Contents

Part I Keynote Talks

Looking at the Interaction Management with New Eyes - Conversational Synchrony and Cooperation using Eye Gaze .................................................. 3
Kristiina Jokinen

Interacting with Purpose (and Feeling!): What Neuropsychology and the Performing Arts Can Tell Us About 'Real' Spoken Language Behaviour ...................................... 5
Roger K. Moore

Part II Speech Recognition and Semantic Analysis

Accessing Web Resources in Different Languages Using a Multilingual Speech Dialog System ................................................................. 9
Hansjörg Hofmann and Andreas Eberhardt and Ute Ehrlich

1 Introduction ................................................................................ 9
2 System Architecture .................................................................... 10
3 Information Extraction from Semi-structured Web Sites ............. 11
4 Generic Speech Dialog ................................................................. 12
5 System Prototype ........................................................................ 12
6 Conclusions ................................................................................ 14
References .................................................................................... 14

New Technique for Handling ASR Errors at the Semantic Level in Spoken Dialogue Systems ......................................................... 17
Ramón López-Cózar, Zoraida Callejas, David Griol and José F. Quesada

1 Introduction ................................................................................ 17
2 Related Work ............................................................................. 18
3 The Proposed Technique ............................................................. 19
3.1 Creation of an initial correction model .................................... 20
3.2 Optimisation of the initial correction model ......................... 20
4 Experiments .............................................................................. 23
| 4.1 Utterance corpora and scenarios | 23 |
| 4.2 Language models for ASR | 25 |
| 4.3 Results | 25 |

5 Conclusions and Future Work | 27

References | 27

**Combining Slot-based Vector Space Model for Voice Book Search** | 31
Cheongjae Lee and Tatsuya Kawahara and Alexander Rudnicky

1 Introduction | 31
2 Related Work | 32
3 Data Collection | 32
3.1 Backend Database | 32
3.2 Query Collection using Amazon Mechanical Turk | 33
4 Book Search Algorithm | 33
4.1 Baseline Vector Space Model (VSM) | 34
4.2 Multiple VSM | 34
4.3 Hybrid VSM and a Back-off Scheme | 36
5 Search Evaluation | 36
5.1 Experiment Set-up | 36
5.2 Evaluation on Textual Queries | 37
5.3 Evaluation on Noisy Queries | 38
6 Conclusion and Discussion | 38
References | 38

**Preprocessing of Dysarthric Speech in Noise Based on CV–Dependent Wiener Filtering** | 41
Ji Hun Park, Woo Kyeong Seong, and Hong Kook Kim

1 Introduction | 41
2 CV–Dependent Wiener Filter | 42
2.1 CV–Classified VAD | 42
2.2 CV–Dependent Wiener Filter | 44
3 Performance Evaluation | 45
4 Conclusion | 46
References | 46

**Conditional Random Fields for Modeling Korean Pronunciation Variation** | 49
Sakriani Sakti, Andrew Finch, Chiori Hori, Hideki Kashioka, Satoshi Nakamura

1 Introduction | 49
2 Speech Recognition Framework | 50
3 Conditional Random Field Approach | 51
4 CRFs Feature Set | 51
5 Experimental Evaluation | 52
6 Conclusions | 54
7 Acknowledgements | 54
An Analysis of the Speech Under Stress Using the Two-Mass Vocal Fold Model
Xiao Yao, Takatoshi Jitsuhiro, Chiyomi Miyajima, Norihide Kitaoka, Kazuya Takeda

1 Introduction .................................................. 58
2 Measuring stress using glottal source ......................... 59
   2.1 Spectral flatness of the glottal flow .................... 59
   2.2 Evaluation of Spectral Flatness Measure ................ 59
3 Simulation using two-mass model .......................... 59
4 Conclusion .................................................. 62
5 Acknowledgements ........................................... 62
References ...................................................... 62

Domain-Adapted Word Segmentation for an Out-of-Domain Language Modeling
Euisok Chung, Hyung-Bae Jeon, Jeon-Gue Park and Yun-Keun Lee
1 Introduction .................................................. 63
2 Domain Adapted Word Segmentation ......................... 64
   2.1 Word Segmentation ..................................... 64
   2.2 Domain Adaptation .................................... 65
   2.3 Unknown Word Extraction ............................. 65
   2.4 Incremental Domain Adaptation ....................... 68
3 Experiments ................................................. 69
   3.1 Word Segmentation Error Reduction .................... 69
   3.2 Incremental Domain Adaptation Experiment .......... 70
4 Discussion ................................................ 71
5 Acknowledgements .......................................... 72
References ...................................................... 72

Part III Multi-Modality for Input and Output

Analysis on Effects of Text-to-Speech and Avatar Agent in Evoking Users’ Spontaneous Listener’s Reactions
Teruhisa Misu, Etsuo Mizukami, Yoshinori Shiga, Shinichi Kawamoto, Hisashi Kawai and Satoshi Nakamura
1 Introduction .................................................. 77
2 Related Works ............................................... 78
3 Construction of Spoken Dialogue TTS ....................... 79
   3.1 Spoken Dialogue Data collection and Model Training . 79
   3.2 Comparison Target .................................... 80
   3.3 Comparison of Prosodic Features of the Synthesized Speech ........................................... 80
4 Construction of Avatar Agent ............................... 81
5 User Experiment ............................................. 82
Development of a Data-driven Framework for Multimodal Interactive Systems
Masahiro Araki and Yuko Mizukami

1 Introduction .......................................................... 91
2 Related Research ......................................................... 92
  2.1 Object-oriented approach for development of spoken dialogue systems .......... 92
  2.2 Data-driven development of Web applications ........................................ 93
  2.3 MMI system architecture .................................................. 93
3 Data-driven framework for MMI system development ....................................... 94
  3.1 Background architecture .................................................. 94
  3.2 Interaction level markup language ............................................ 94
  3.3 Dialog flow description ..................................................... 95
4 Object-oriented modeling language .................................................................. 97
  4.1 Language specification .......................................................... 97
  4.2 Rapid initial prototyping ......................................................... 98
  4.3 Extension ........................................................................... 99
5Conclusion and Future Research ..................................................................... 100
References ......................................................................................... 100
Contents

2.1 Communication Situations ........................................ 115
2.2 Data Acquisition ................................................. 116
3 System Overview .................................................. 117
  3.1 HMM Training on Situation Dependent Speech Corpora .... 117
  3.2 System Configuration ......................................... 118
4 Evaluation .......................................................... 119
  4.1 Experimental Setup .......................................... 119
  4.2 Experimental Result .......................................... 120
5 Conclusion .......................................................... 121
References ............................................................ 122

An Event-Based Conversational System for the Nao Robot .......... 125
Ivana Kruijff-Korbayova, Georgios Athanasopoulos, Aryel Beck, Piero Cosi, Heriberto Cuayáhuítl, Tomas Dekens, Valentin Enescu, Antoine Hiolle, Bernd Kiefer, Hichem Sahli, Marc Schröder, Giacomo Sommavilla, Fabio Tesser and Werner Verhelst
1 Introduction ......................................................... 125
2 Event-Based Component Integration ................................ 126
3 The Integrated System ............................................. 126
  3.1 Dialogue Manager (DM) ...................................... 127
  3.2 Audio Front End (AFE) and Voice Activity Detection (VAD) ......................................................... 127
  3.3 Automatic Speech Recognition (ASR) ....................... 127
  3.4 Natural Language Understanding (NLU) .................... 128
  3.5 Natural Language Generation (NLG) ....................... 128
  3.6 Text-To-Speech Synthesis (TTS) ............................ 129
  3.7 Gesture Recognition and Understanding (GRU) .......... 129
  3.8 Non-Verbal Behavior Planning (NVBP) & Motor Control (MC) ......................................................... 130
4 Experience from Experiments and Conclusions .................. 130
References ............................................................ 131

Towards Learning Human-Robot Dialogue Policies Combining Speech and Visual Beliefs ........................................ 133
Heriberto Cuayáhuítl, Ivana Kruijff-Korbayová
1 Introduction ......................................................... 133
2 Learning Human-Robot Dialogues Under Uncertainty .......... 134
3 Using Bayesian-Relational State Representations for Optimizing Human-Robot Dialogues ........................................ 134
4 Experiments and Results .......................................... 135
  4.1 The Simulated Conversational Environment ................ 135
  4.2 Characterization of the Learning Agent .................... 136
  4.3 Experimental Results ........................................ 138
5 Conclusion and Future Work ...................................... 139
References ............................................................ 139
Part IV User Modelling

JAM: Java-based Associative Memory ........................................ 143
Robert Pröpper, Felix Putze, Tanja Schultz
1 Introduction ................................................................. 143
2 Related Work ............................................................... 144
3 Architecture ................................................................. 146
  3.1 Knowledge Structure ................................................. 146
  3.2 Memory Dynamics ................................................... 147
4 Implementation ............................................................ 148
5 Evaluation ................................................................. 150
  5.1 Survey ................................................................. 150
  5.2 Conversation ........................................................ 152
References ................................................................. 154

Conversation Peculiarities of People with Different Verbal Intelligence . . 157
Kseniya Zablotskay, Umair Rahim, Sergey Zablotskiy, Steffen Walter, Wolfgang Minker
1 Introduction ................................................................. 158
2 Method ................................................................. 158
  2.1 Corpus Collection .................................................... 158
  2.2 Feature Extraction ................................................... 159
3 Experiments and Results ................................................ 160
4 Discussions and Future Work ........................................... 161
References ................................................................. 162

Merging Intention and Emotion to Develop Adaptive Dialogue Systems . . 165
Zoraida Callejas, David Griol, Ramón López-Cózar, Gonzalo Espejo, Nieves Ábalos
1 Introduction and related work ........................................... 165
2 Our proposal ............................................................. 166
  2.1 The emotion recognizer ............................................. 166
  2.2 The intention recognizer ........................................... 167
3 The enhanced UAH dialogue system .................................. 168
4 Experiments ............................................................. 170
5 Conclusions and future work .......................................... 173
6 Acknowledgments ....................................................... 174
References ................................................................. 174

All Users Are (Not) Equal - The Influence of User Characteristics on Perceived Quality, Modality Choice and Performance ........................................ 175
Ina Wechsung, Matthias Schulz, Klaus-Peter Engelbrecht, Julia Niemann and Sebastian Möller
1 Introduction ............................................................. 175
2 Related Work ........................................................... 176
3 Method ................................................................. 177
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Experimental Set-Up</td>
<td>178</td>
</tr>
<tr>
<td>3.2 Measures</td>
<td>178</td>
</tr>
<tr>
<td>3.3 Procedure</td>
<td>179</td>
</tr>
<tr>
<td>4 Results</td>
<td>180</td>
</tr>
<tr>
<td>4.1 Factors Influencing Performance</td>
<td>180</td>
</tr>
<tr>
<td>4.2 Factors Influencing Modality Choice</td>
<td>181</td>
</tr>
<tr>
<td>4.3 Factors Influencing Quality Perceptions</td>
<td>182</td>
</tr>
<tr>
<td>5 Discussion and Conclusion</td>
<td>184</td>
</tr>
</tbody>
</table>

## Part V Dialogue Management

### Parallel Computing and Practical Constraints when applying the Standard POMDP Belief Update Formalism to Spoken Dialogue Management

Paul A. Crook, Brieuc Roblin, Hans-Wolfgang Loidl and Oliver Lemon

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>189</td>
</tr>
<tr>
<td>1.1 Paper Structure</td>
<td>190</td>
</tr>
<tr>
<td>2 Background</td>
<td>191</td>
</tr>
<tr>
<td>2.1 Typical SDS</td>
<td>191</td>
</tr>
<tr>
<td>2.2 POMDP Formalism</td>
<td>191</td>
</tr>
<tr>
<td>3 Related Literature</td>
<td>192</td>
</tr>
<tr>
<td>3.1 Dialogue Response Time</td>
<td>192</td>
</tr>
<tr>
<td>4 POMDP DM</td>
<td>193</td>
</tr>
<tr>
<td>4.1 Fixed Users’ Goals during Dialogues</td>
<td>195</td>
</tr>
<tr>
<td>5 Methodology</td>
<td>195</td>
</tr>
<tr>
<td>6 Dense POMDP Belief Updates</td>
<td>196</td>
</tr>
<tr>
<td>7 Limits for SDS DM</td>
<td>197</td>
</tr>
<tr>
<td>7.1 Practical Systems</td>
<td>198</td>
</tr>
<tr>
<td>8 Conclusions</td>
<td>198</td>
</tr>
</tbody>
</table>

### Ranking Dialog Acts using Discourse Coherence Indicator for Language Tutoring Dialog Systems

Hyungjong Noh, Sungjin Lee, Kyungduk Kim, Kyusong Lee, Gary Geunbae Lee

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>204</td>
</tr>
<tr>
<td>2 Related Work</td>
<td>205</td>
</tr>
<tr>
<td>3 Discourse Coherence Indicator for Dialog Acts</td>
<td>206</td>
</tr>
<tr>
<td>3.1 Necessity of Discourse Coherence Indicator</td>
<td>206</td>
</tr>
<tr>
<td>3.2 Formulation</td>
<td>207</td>
</tr>
<tr>
<td>4 Similarity between discourse histories</td>
<td>209</td>
</tr>
<tr>
<td>4.1 Enhanced Levenshtein Distance</td>
<td>209</td>
</tr>
<tr>
<td>4.2 Using DCI</td>
<td>209</td>
</tr>
<tr>
<td>4.3 Using Discount Rate Parameter</td>
<td>210</td>
</tr>
<tr>
<td>4.4 Ranking Score Normalization</td>
<td>210</td>
</tr>
</tbody>
</table>
On-line detection of task incompletion for spoken dialog systems using utterance and behavior tag N-gram vectors

Sunao Hara, Norihide Kitaoka, and Kazuya Takeda

1 Introduction ................................................................. 215
2 Spoken dialog corpus of a music retrieval task ....................... 217
3 Feature construction from dialogs .................................... 218
3.1 Encoding utterances and behaviors as tags ......................... 218
3.2 Construction of tag N-gram feature ................................. 218
3.3 Construction of interaction parameter features .................... 219
3.4 Training classifiers based on SVM ................................. 220
4 Detection of task-incomplete dialogs .................................. 220
4.1 Evaluation of off-line detection ..................................... 220
4.2 Evaluation of on-line detection performance ........................ 222
5 Conclusion ..................................................................... 223
References ..................................................................... 224

Integration of Statistical Dialog Management Techniques to Implement Commercial Dialog Systems ........................................ 227

David Griol, Zoraida Callejas, Ramón López-Cózar

1 Introduction ..................................................................... 227
2 Our Proposal to Introduce Statistical Methodologies in Commercial Dialog Systems ..................................... 229
2.1 Implementation by means of the Standard VoiceXML ............ 231
2.2 User Simulation to Learn the Dialog Model .......................... 232
3 Development of a Railway Information System using the Proposed Technique ......................................... 233
4 Evaluation of the Developed Dialog System .......................... 235
5 Conclusions and Future Work ............................................. 237
References ..................................................................... 237

A Theoretical Framework for a User-Centered Spoken Dialog Manager .......................................................... 241

Stefan Ultes, Tobias Heinroth, Alexander Schmitt, Wolfgang Minker

1 Introduction ..................................................................... 241
2 Related Work ..................................................................... 242
3 Motivation ..................................................................... 243
4 Theoretical Framework ..................................................... 244
5 Conclusion ..................................................................... 245
6 Acknowledgement ........................................................... 245
References ..................................................................... 246
## Contents

### Using probabilistic logic for dialogue strategy selection
Ian O’Neill, Philip Hanna, Anbu Yue and Weiru Liu

1. Adaptive Dialogue ........................................ 247
2. The Experiment and its Evaluation ................. 249
3. Conclusions ............................................. 252
References .................................................. 252

### Starting to Cook a Coaching Dialogue System in the Olympus framework
Joana Paulo Pardal and Nuno J. Mamede

1. Introduction ........................................... 255
2. COOKCOACH ........................................ 256
   2.1 OLYMPUS/RAVENCLAW framework .......... 257
   2.2 Interaction design ............................... 259
   2.3 Recipes Model and OntoChef ............... 260
   2.4 Acquiring recipes ............................ 261
   2.5 Interface Design ............................... 263
   2.6 Cook Tutor ...................................... 263
3. Pursuing ontology-based dialogue systems .......... 263
   3.1 Reasoning ....................................... 264
   3.2 Learning ......................................... 264
   3.3 New Systems ..................................... 265
4. Conclusion ............................................. 265
References .................................................. 266

### Part VI Evaluation Strategies and Paradigms

#### Performance of an Ad-hoc User Simulation in a Formative Evaluation of a Spoken Dialog System
Klaus-Peter Engelbrecht, Stefan Schmidt, Sebastian Möller

1. Introduction ........................................... 271
2. Experimental Data .................................... 273
3. User and Speech Understanding Models .......... 274
4. Creating a List of Usability Problems from Real User Data .... 275
5. Problem Discovery in the Simulated Corpora .... 277
6. Preparation of Data for Log File Inspection .... 279
7. Discussion ............................................ 280
8. Conclusions and Future Work .................... 282
References .................................................. 282

#### Adapting Dialogue to User Emotion - A Wizard-of-Oz study for adaptation strategies
Gregor Bertrand, Florian Nothdurft, Wolfgang Minker, Harald Traue and Steffen Walter

1. Introduction ........................................... 286
2. Related Work ......................................... 286
3. Our Experiment ...................................... 287
How context determines perceived quality and modality choice. Secondary task paradigm applied to the evaluation of multimodal interfaces. Ina Wechsung, Robert Schleicher and Sebastian Möller

Part VII Prototypes and Products


A Dialogue System for Conversational NPCs Tina Klüwer, Peter Adolphs, Feiyu Xu and Hans Uszkoreit

Embedded Conversational Engine for Natural Language Interaction in Spanish Marcos Santos-Pérez, Eva González-Parada and José Manuel Cano-García
Adding Speech to a Robotics Simulator ................................. 375
Graham Wilcock and Kristiina Jokinen
1 Introduction ......................................................... 375
2 Pyro Robotics .................................................. 377
3 Spoken Interaction .............................................. 378
4 Spoken Dialogues .............................................. 379
5 Further Work .................................................... 380
References ......................................................... 380

Index ............................................................................. 381
List of Contributors

Nieves Ábalos
Dept. of Languages and Computer Systems, CITIC-UGR, University of Granada, Spain, e-mail: nayade@correo.ugr.es

Peter Adolphs
DFKI, Berlin, Germany, e-mail: peter.adolphs@dfki.de

Masahiro Araki
Kyoto Institute of Technology, Kyoto, Japan, e-mail: araki@kit.jp

Georgios Athanasopoulos
IBBT, Vrije Universiteit Brussel, Dept. ETRO-DSSP, Belgium, e-mail: gathanas@etro.vub.ac.be

Aryel Beck
School of Computer Science, University of Hertfordshire, UK, e-mail: arylenebeck@gmail.com

Gregor Bertrand
Institute of Information Technology, University of Ulm, Germany, e-mail: gregor.bertrand@uni-ulm.de

Zoraida Callejas
Dept. of Languages and Computer Systems, CITIC-UGR, University of Granada, Spain, e-mail: zoraida@ugr.es

José Manuel Cano-García
Electronic Technology Department, University of Málaga, Spain, e-mail: jcgarcia@uma.es

Euisok Chung
Speech Processing Team, ETRI, Daejeon, Korea, e-mail: eschung@etri.re.kr
Piero Cosi
Istituto di Scienze e Tecnologie della Cognizione, ISTC, C.N.R., Italy, e-mail: piero.cosi@pd.istc.cnr.it

Paul A. Crook
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK, e-mail: p.a.crook@hw.ac.uk

Heriberto Cuayahuitl
DFKI GmbH, Language Technology Lab, Saarbrücken, Germany, e-mail: h.cuayahuitl@gmail.com

Tomas Dekens
IBBT, Vrije Universiteit Brussel, Dept. ETRO-DSSP, Belgium, e-mail: tdekens@etro.vub.ac.be

Andreas Eberhardt
BitTwister IT GmbH, Senden, Germany, e-mail: andreas.eberhardt@bittwister.com

Ute Ehrlich
Daimler AG, Ulm, Germany, e-mail: ute.ehrlich@daimler.com

Valentin Enescu
IBBT, Vrije Universiteit Brussel, Dept. ETRO-DSSP, Belgium, e-mail: venescu@etro.vub.ac.be

Klaus-Peter Engelbrecht
Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, Germany, e-mail: klaus-peter.engelbrecht@telekom.de

Gonzalo Espejo
Dept. of Languages and Computer Systems, CITIC-UGR, University of Granada, Spain, e-mail: gonzalez@uma.es

Andrew Finch
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: andrew.finch@nict.go.jp

Shinya Fujie
Waseda Institute of Advanced Study, Japan, e-mail: fujie@pcl.cs.waseda.ac.jp

Eva González-Parada
Electronic Technology Department, University of Málaga, Spain, e-mail: gonzalez@uma.es

David Griol
Dept. of Computer Science, Carlos III University of Madrid, Spain, e-mail: dgriol@inf.uc3m.es
Philip Hanna  
School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, Northern Ireland, e-mail: p.hanna@qub.ac.uk

Sunao Hara  
Graduate School of Information Science, Nagoya University, Aichi, Japan, e-mail: naoh@nagoya-u.jp

Tobias Heinroth  
Institute of Information Technology, University of Ulm, Germany, e-mail: tobias.heinroth@uni-ulm.de

Antoine Hiolle  
School of Computer Science, University of Hertfordshire, UK, e-mail: antoine.hiolle@gmail.com

Hansjörg Hofmann  
Daimler AG, Ulm, Germany, e-mail: hansjoerg.hofmann@daimler.com

Chiori Hori  
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: chiori.hori@nict.go.jp

Kazuhiro Iwata  
Waseda University, Tokyo, Japan, e-mail: iwata@pci.cs.waseda.ac.jp

Hyung-Bae Jeon  
Speech Processing Team, ETRI, Daejeon, Korea, e-mail: hbjeon@etri.re.kr

Takatoshi Jitsuhiro  
Department of Media Informatics, Aichi University of Technology, Gamagori Japan e-mail: jitsuhiro@aut.ac.jp

Kristiina Jokinen  
University of Helsinki, Finland, e-mail: kristiina.jokinen@helsinki.fi

Hideki Kashioka  
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: hideki.kashioka@nict.go.jp

Tatsuya Kawahara  
Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan, e-mail: kawahara@ar.media.kyoto-u.ac.jp

Hisashi Kawai  
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: hisashi.kawai@nict.go.jp

Shinichi Kawamoto  
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: shinichi.kawamoto@nict.go.jp
List of Contributors

Bernd Kiefer
DFKI GmbH, Language Technology Lab, Saarbrücken, Germany, e-mail: kiefer@dfki.de

Hong Kook Kim
School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea, e-mail: hongkook@gist.ac.kr

Kyungduk Kim
POSTECH, San 31, Hyoja-Dong, Pohang, 790-784, South Korea, e-mail: getta@postech.ac.kr

Norihide Kitaoka
Graduate School of Information Science, Nagoya University, Aichi, Japan, e-mail: kitaoka@nagoya-u.jp

Tina Klüwer
DFKI, Berlin, Germany, e-mail: tina.kluewer@dfki.de

Tetsunori Kobayashi
Department of Computer Science and Engineering, Waseda University, Japan, e-mail: koba@waseda.jp

Kazunori Komatani
Graduate School of Engineering, Nagoya University, Japan, e-mail: komatani@nuee.nagoya-u.ac.jp

Ivana Kruijff-Korbayová
DFKI GmbH, Language Technology Lab, Saarbrücken, Germany, e-mail: ivana.kruijff@dfki.de

Cheongjae Lee
Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan, e-mail: lcj80@ar.media.kyoto-u.ac.jp

Gary Geunbae Lee
POSTECH, San 31, Hyoja-Dong, Pohang, 790-784, South Korea, e-mail: gblee@postech.ac.kr

Kyusong Lee
POSTECH, San 31, Hyoja-Dong, Pohang, 790-784, South Korea, e-mail: kyusonglee@postech.ac.kr

Sungjin Lee
POSTECH, San 31, Hyoja-Dong, Pohang, 790-784, South Korea, e-mail: junion@postech.ac.kr

Yun-Keun Lee
Speech Processing Team, ETRI, Daejeon, Korea, e-mail: yklee@etri.re.kr
List of Contributors

Oliver Lemon
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK, e-mail: o.lemon@hw.ac.uk

Weiru Liu
School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, Northern Ireland, e-mail: w.liu@qub.ac.uk

Hans-Wolfgang Loidl
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK, e-mail: h.w.loidl@hw.ac.uk

Ramón López-Cózar
Dept. of Languages and Computer Systems, CITIC-UGR, University of Granada, Spain, e-mail: rlopezc@ugr.es

Nuno J. Mamede
Spoken Language Systems Laboratory, L2F – INESC-ID and IST, Technical University of Lisbon, Portugal, e-mail: nuno.mamede@l2f.inesc-id.pt

Kyoko Matsuyama
Graduate School of Informatics, Kyoto University, Japan, e-mail: matuyama@kuis.kyoto-u.ac.jp

Yoichi Matsuyama
Department of Computer Science and Engineering, Waseda University, Japan, e-mail: matsuyama@pcl.cs.waseda.ac.jp

Wolfgang Minker
Institute of Information Technology, University of Ulm, Germany, e-mail: wolfgang.minker@uni-ulm.de

Teruhisa Misu
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: teruhisa.misu@nict.go.jp

Chiyomi Miyajima
Graduate School of Information Science, Nagoya University, Aichi, Japan, e-mail: miyajima@nagoya-u.jp

Etsuo Mizukami
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: etsuo.mizukami@nict.go.jp

Yuko Mizukami
Kyoto Institute of Technology, Kyoto, Japan, e-mail: m9622034@edu.kit.ac.jp

Sebastian Möller
Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, Germany, e-mail: sebastian.moeller@telekom.de
Roger K. Moore  
Department of Computer Science, University of Sheffield, UK, e-mail: 
*r.k.moore@dcs.shef.ac.uk*

Satoshi Nakamura  
Nara Institute of Science and Technology (NAIST), Nara, Japan, e-mail:  
*s-nakamura@is.naist.jp*

Julia Niemann  
Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, Germany,  
e-mail: *julia.niemann@telekom.de*

Hyungjong Noh  
POSTECH, San 31, Hyoja-Dong, Pohang, 790-784, South Korea, e-mail:  
*nohhj@postech.ac.kr*

Florian Nothdurft  
Institute of Information Technology, University of Ulm, Germany, e-mail:  
*florian.nothdurft@uni-ulm.de*

Tetsuya Ogata  
Graduate School of Informatics, Kyoto University, Japan, e-mail:  
*ogata@kuis.kyoto-u.ac.jp*

Hiroshi G. Okuno  
Graduate School of Informatics, Kyoto University, Japan, e-mail:  
*okuno@kuis.kyoto-u.ac.jp*

Ian O’Neill  
School of Electronics, Electrical Engineering and Computer Science, Queen’s  
University Belfast, Northern Ireland, e-mail: *i.oneill@qub.ac.uk*

Joana Paulo Pardal  
Spoken Language Systems Laboratory, L²F – INESC-ID and IST, Technical University of Lisbon, Portugal, e-mail:  
*joana.paulo.pardal@l2f.inesc-id.pt*

Jeon-Gue Park  
Speech Processing Team, ETRI, Daejeon, Korea, e-mail: *jgp@etri.re.kr*

Ji Hun Park  
School of Information and Communications, Gwangju Institute of Science and  
Technology, Gwangju 500-712, Korea, e-mail: *jh_park@gist.ac.kr*

Robert Pröpper  
Cognitive Systems Lab (CSL), Karlsruhe Institute of Technology (KIT), Germany,  
e-mail: *robert.proepper@student.kit.edu*

Felix Putze  
Cognitive Systems Lab (CSL), Karlsruhe Institute of Technology (KIT), Germany,  
e-mail: *felix.putze@kit.edu*
José F. Quesada  
Dept. of Computer Science and Artificial Intelligence, University of Seville, Spain,  
e-mail: Jose.Quesada@infinity.es

Umair Rahim  
Institute of Information Technology, University of Ulm, Germany, e-mail: umair.rahim@uni-ulm.de

Norbert Reithinger  
Deutsches Forschungszentrum für Künstliche Intelligenz, Projektbüro Berlin,  
Germany, e-mail: norbert.reithinger@dfki.de

Brieuc Roblin  
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK, e-mail: brieuc.roblin@gmail.com

Roland Roller  
Deutsches Forschungszentrum für Künstliche Intelligenz, Projektbüro Berlin,  
Germany, e-mail: roland.roller@dfki.de

Alexander Rudnicky  
Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA,  
e-mail: alex@cs.cmu.edu

Hichem Sahli  
IBBT, Vrije Universiteit Brussel, Dept. ETRO-DSSP, Belgium, e-mail: hicem.sahli@etro.vub.ac.be

Akihiro Saito  
Department of Computer Science and Engineering, Waseda University, Japan,  
e-mail: saito@pcl.cs.waseda.ac.jp

Sakriani Sakti  
Nara Institute of Science and Technology (NAIST), Nara, Japan, e-mail: ssakti@is.naist.jp

Marcos Santos-Pérez  
Electronic Technology Department, University of Málaga, Spain, e-mail: marcos_sape@uma.es

Tatjana Scheffler  
Deutsches Forschungszentrum für Künstliche Intelligenz, Projektbüro Berlin,  
Germany, e-mail: tatjana.scheffler@dfki.de

Robert Schleicher  
Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, Germany,  
e-mail: robert.schleicher@telekom.de

Stefan Schmidt  
Quality and Usability Lab, Deutsche Telekom Laboratories, TU-Berlin, Germany  
e-mail: Stefan.Schmidt01@Telekom.de
List of Contributors

Alexander Schmitt
Institute of Information Technology, University of Ulm, Germany, e-mail: alexander.schmitt@uni-ulm.de

Marc Schröder
DFKI GmbH, Language Technology Lab, Saarbrücken, Germany, e-mail: marc.schroeder@dfki.de

Tanja Schultz
Cognitive Systems Lab (CSL), Karlsruhe Institute of Technology (KIT), Germany, e-mail: tanja.schultz@kit.edu

Matthias Schulz
Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, Germany, e-mail: matthias-schulz@telekom.de

Woo Kyeong Seong
School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea, e-mail: wkseong@gist.ac.kr

Yoshinori Shiga
National Institute of Information and Communications Technology (NICT), Kyoto, Japan, e-mail: yoshinori.shiga@nict.go.jp

Giacomo Sommavilla
Istituto di Scienze e Tecnologie della Cognizione, ISTC, C.N.R., Italy, e-mail: giacomo.sommavilla@pd.istc.cnr.it

Kazuya Takeda
Graduate School of Information Science, Nagoya University, Aichi, Japan, e-mail: kazuya.takeda@nagoya-u.jp

Ryu Takeda
Graduate School of Informatics, Kyoto University, Japan, e-mail: rtakeda@kuis.kyoto-u.ac.jp

Fabio Tesser
Istituto di Scienze e Tecnologie della Cognizione, ISTC, C.N.R., Italy, e-mail: fabio.tesser@gmail.com

Harald C. Traue
Medical Psychology, University Clinic for Psychosomatic Medicine and Psychotherapy, Ulm University, Germany, e-mail: harald.traue@uni-ulm.de

Stefan Ultes
Institute of Information Technology, University of Ulm, Germany, e-mail: stefan.ultes@uni-ulm.de

Hans Uszkoreit
DFKI, Berlin, Germany, e-mail: uszkoreit@dfki.de
Werner Verhelst  
IBBT, Vrije Universiteit Brussel, Dept. ETRO-DSSP, Belgium, e-mail: wverhelst@etro.vub.ac.be

Steffen Walter  
Medical Psychology, University Clinic for Psychosomatic Medicine and Psychotherapy, Ulm University, Germany,  
e-mail: steffen.walter@uni-ulm.de

Ina Wechsung  
Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, Germany,  
e-mail: ina.wechsung@telekom.de

Graham Wilcock  
University of Helsinki, Finland, e-mail: graham.wilcock@helsinki.fi

Feiyu Xu  
DFKI, Berlin, Germany, e-mail: feiyu.xu@dfki.de

Yushi Xu  
MIT Computer Science and Artificial Intelligence Laboratory, USA, e-mail: yushixu@csail.mit.edu

Xiao Yao  
Graduate School of Information Science, Nagoya University, Aichi, Japan, e-mail: xiao.yao@g.sp.m.is.nagoya-u.ac.jp

Anbu Yue  
School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, Northern Ireland, e-mail: a.yue@qub.ac.uk

Kseniya Zablotskaya  
Institute of Information Technology, University of Ulm, Germany, e-mail: kseniya.zablotskaya@uni-ulm.de

Sergey Zablotskiy  
Institute of Information Technology, University of Ulm, Germany, e-mail: sergey.zablotskiy@uni-ulm.de
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFE</td>
<td>Audio Front End</td>
</tr>
<tr>
<td>AIML</td>
<td>Artificial Intelligence Mark-up Language</td>
</tr>
<tr>
<td>ALICE</td>
<td>Artificial Linguistic Internet Computer Entity</td>
</tr>
<tr>
<td>AmI</td>
<td>Ambient Intelligence</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Simple One-Way Analysis of Variance</td>
</tr>
<tr>
<td>AODE</td>
<td>Averaged One-Dependence Estimators</td>
</tr>
<tr>
<td>AOST</td>
<td>Automatic Orthographic Transcriber</td>
</tr>
<tr>
<td>ASR</td>
<td>Automatic Speech Recognition</td>
</tr>
<tr>
<td>AST</td>
<td>Automatic Semantic Transcriber</td>
</tr>
<tr>
<td>AVP</td>
<td>Attribute-Value-Pair</td>
</tr>
<tr>
<td>BAS</td>
<td>Behavioral Approach System</td>
</tr>
<tr>
<td>BC</td>
<td>Back Channel</td>
</tr>
<tr>
<td>BFI-S</td>
<td>Big Five Inventory-Short</td>
</tr>
<tr>
<td>BIS</td>
<td>Behavioral Inhibition System</td>
</tr>
<tr>
<td>CCG</td>
<td>Combinatory Categorial Grammar</td>
</tr>
<tr>
<td>CER</td>
<td>Concept Error Rate</td>
</tr>
<tr>
<td>CM</td>
<td>Correction Model</td>
</tr>
<tr>
<td>CRF</td>
<td>Conditional Random Field</td>
</tr>
<tr>
<td>CV</td>
<td>Consonant-Vowel</td>
</tr>
<tr>
<td>CVC</td>
<td>Consonant-Vowel-Consonant</td>
</tr>
<tr>
<td>DAMSL</td>
<td>Dialog Act Markup in Several Layers</td>
</tr>
<tr>
<td>DCI</td>
<td>Discourse Coherency Indicator</td>
</tr>
<tr>
<td>DD</td>
<td>Decision-Directed</td>
</tr>
<tr>
<td>DM</td>
<td>Dialogue Manager</td>
</tr>
<tr>
<td>DR</td>
<td>Dialogue Register</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyography</td>
</tr>
<tr>
<td>ERQ</td>
<td>Emotion Regulation Questionnaire</td>
</tr>
<tr>
<td>ERR</td>
<td>Error Reduction Rate</td>
</tr>
<tr>
<td>FLOPS</td>
<td>Floating Point Operations Per Second</td>
</tr>
<tr>
<td>FOD</td>
<td>First-Order Difference</td>
</tr>
<tr>
<td>FSM</td>
<td>Finite State Machine</td>
</tr>
<tr>
<td>Acronyms</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>GPGPU</td>
<td>General Purpose Graphics Processing Unit</td>
</tr>
<tr>
<td>GRU</td>
<td>Gesture Recognition and Understanding</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HAWIE</td>
<td>Hamburg Wechsler Intelligence Test for Adults</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction</td>
</tr>
<tr>
<td>HDD</td>
<td>Hard Disk Drive</td>
</tr>
<tr>
<td>HIS</td>
<td>Hidden Information State</td>
</tr>
<tr>
<td>HIT</td>
<td>Human Intelligence Task</td>
</tr>
<tr>
<td>HMM</td>
<td>Hidden Markov Model</td>
</tr>
<tr>
<td>HSV</td>
<td>Hue Lightness Saturation Model</td>
</tr>
<tr>
<td>HVSM</td>
<td>Hybrid Vector Space Model</td>
</tr>
<tr>
<td>IBBT</td>
<td>Interdisciplinary institute for BroadBand Technology</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>IE</td>
<td>Information Extraction</td>
</tr>
<tr>
<td>IPA</td>
<td>Intelligent Procedure Assistant</td>
</tr>
<tr>
<td>IPTV</td>
<td>Internet-based Television</td>
</tr>
<tr>
<td>IPU</td>
<td>Inter-Pausal Unit</td>
</tr>
<tr>
<td>IR</td>
<td>Implicit Recovery</td>
</tr>
<tr>
<td>IRS</td>
<td>Information Retrieval Systems</td>
</tr>
<tr>
<td>ISS</td>
<td>International Space Station</td>
</tr>
<tr>
<td>ISTC</td>
<td>Institute of Cognitive Sciences and Technologies</td>
</tr>
<tr>
<td>ITSCJ</td>
<td>Information Technology Standards Commission of Japan</td>
</tr>
<tr>
<td>ITU</td>
<td>International Telecommunication Union</td>
</tr>
<tr>
<td>ITU-T</td>
<td>Telecommunication Standardization Sector</td>
</tr>
<tr>
<td>IVR</td>
<td>Interactive Voice Response</td>
</tr>
<tr>
<td>JSGF</td>
<td>Java Speech Grammar Format</td>
</tr>
<tr>
<td>KMS</td>
<td>Known-word Mono-Syllable</td>
</tr>
<tr>
<td>KWA</td>
<td>KeyWord Accuracy</td>
</tr>
<tr>
<td>LARRI</td>
<td>Language-based Agent for Retrieval of Repair Information</td>
</tr>
<tr>
<td>LIWC</td>
<td>Linguistic Inquiry and Word Count</td>
</tr>
<tr>
<td>LP</td>
<td>Linear Prediction</td>
</tr>
<tr>
<td>LRT</td>
<td>Likelihood Ratio Test</td>
</tr>
<tr>
<td>LVCSR</td>
<td>Large Vocabulary Continuous Speech Recognition</td>
</tr>
<tr>
<td>MARY</td>
<td>Modular Architecture for Research on speech Synthesis</td>
</tr>
<tr>
<td>MC</td>
<td>Motor Control</td>
</tr>
<tr>
<td>MDL</td>
<td>Minimum Description Length</td>
</tr>
<tr>
<td>MDP</td>
<td>Markov Decision Process</td>
</tr>
<tr>
<td>MEP</td>
<td>Maximum Entropy Probability</td>
</tr>
<tr>
<td>MFCC</td>
<td>Mel-Frequency Cepstral Coefficients</td>
</tr>
<tr>
<td>MILM</td>
<td>Multimodal Interaction Markup Language</td>
</tr>
<tr>
<td>MMG</td>
<td>Multi Motive Grid</td>
</tr>
<tr>
<td>MMI</td>
<td>MultiModal Interaction</td>
</tr>
<tr>
<td>MOS</td>
<td>Mean Opinion Score</td>
</tr>
<tr>
<td>MRA</td>
<td>Multiple Regression Analysis</td>
</tr>
<tr>
<td>MRDA</td>
<td>Meeting Recorder Dialog Act</td>
</tr>
</tbody>
</table>
MRR  Mean Reciprocal Rank
MRT  Multiple Resource Theory
MSE  Mean Square Error
MT  Machine Translation
MTurk  Amazon Mechanical Turk
MVC  Model View Controller
MVSM  Multiple Vector Space Model
NLG  Natural Language Generation
NLU  Natural Language Understanding
NPC  Non Player Character
NVBP  Non-verbal Behavior Planning
ODB  Object-oriented DataBase
ODP  Ontology-based Dialogue Platform
OOG  Out-Of-Grammar
OOV  Out-Of-Vocabulary
OWL  Web Ontology Language
PDL  Partially Observable Markov Decision Process
POS  Part-Of-Speech
PSAT  Probabilistic Satisfiability
PSTN  Public Switched Telephone Network
RDB  Relational DataBase
RDF  Resource Description Framework
RL  Reinforcement Learning
ROC  Receiver Operating Characteristic
SAT  Satisfiability
SDC  Spoken Dialog Challenge
SDM  Spoken Dialogue Manager
SDS  Spoken Dialogue System
SEA-scale  Subjektiv Erlebte Anstrengung (Subjectively Perceived Effort) Scale
SEM  Stochastic Expectation-Maximization (algorithm)
SIMD  Single Instruction, Multiple Data
SLDS  Spoken Language Dialogue System
SLU  Spoken Language Understanding
SNR  Signal-to-Noise Ratio
SR  Sentence Recognition
SRGS  Speech Recognition Grammar Specification
SSE  Streaming SIMD Extension
SSS  Successive State Splitting
STRAIGHT  Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrogram
SU  Sentence Understanding
SVM  Support Vector Machine
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVSM</td>
<td>Single Vector Space Model</td>
</tr>
<tr>
<td>S.ERR</td>
<td>Sentence Error Rate</td>
</tr>
<tr>
<td>TALK</td>
<td>Tools for Ambient Linguistic Knowledge</td>
</tr>
<tr>
<td>TC</td>
<td>Task Completion</td>
</tr>
<tr>
<td>TTS</td>
<td>Text-To-Speech (synthesis)</td>
</tr>
<tr>
<td>UAH</td>
<td>Universidad al Habla (University on the Line)</td>
</tr>
<tr>
<td>URBI</td>
<td>Universal Robot Body Interface</td>
</tr>
<tr>
<td>UMS</td>
<td>Unknown-word Mono-Syllable</td>
</tr>
<tr>
<td>UW</td>
<td>Unknown Word</td>
</tr>
<tr>
<td>VAD</td>
<td>Voice Activity Detection</td>
</tr>
<tr>
<td>VoiceXML</td>
<td>Voice Extensible Markup Language</td>
</tr>
<tr>
<td>VOP</td>
<td>Voice Onset Point</td>
</tr>
<tr>
<td>VSM</td>
<td>Vector Space Model</td>
</tr>
<tr>
<td>VUB</td>
<td>Free University of Brussels</td>
</tr>
<tr>
<td>VXML</td>
<td>VoiceXML</td>
</tr>
<tr>
<td>WA</td>
<td>Word Accuracy</td>
</tr>
<tr>
<td>WER</td>
<td>Word Error Rate</td>
</tr>
<tr>
<td>WFST</td>
<td>Weighted Finite State Transducer</td>
</tr>
<tr>
<td>WoZ</td>
<td>Wizard-of-Oz</td>
</tr>
<tr>
<td>ZODB</td>
<td>Zope Object-oriented DataBase</td>
</tr>
</tbody>
</table>