

# **2006 IEEE Spoken Language Technology Workshop**

**Palm Beach, Aruba  
10-13 December 2006**



**IEEE Catalog Number:**  
**ISBN:**

**06EX1646**  
**1-4244-0872-5**

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. Other copying, reprint, or reproduction requests should be addressed to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved. Copyright © 2006 by the Institute of Electrical and Electronics Engineers, Inc.

*The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Signal Processing Society or the Institute of Electrical and Electronics Engineers, Inc.*

IEEE Catalog Number 06EX1646

ISBN 1-4244-0872-5

Library of Congress: 2006938742

Produced by Conference Management Services, Inc.



# TABLE OF CONTENTS

---

## **KEY-M: MONDAY KEYNOTE**

<b>KEY-M.1: USING POMDPS FOR DIALOG MANAGEMENT .....</b>	<b>8</b>
<i>Steve Young, Cambridge University, United Kingdom</i>	

## **KEY-T: TUESDAY KEYNOTE**

<b>KEY-T.1: SPEECH TECHNOLOGY OPPORTUNITIES AND CHALLENGES.....</b>	<b>1</b>
<i>David Nahamoo, Speech CTO, IBM Research, United States</i>	

## **KEY-W: WEDNESDAY KEYNOTE**

<b>KEY-W.1: NO MORE STRINGS, PLEASE.....</b>	<b>2</b>
<i>Kevin Knight, USC/Information Sciences Institute, United States</i>	

## **INV-M: MONDAY INVITED TUTORIALS**

<b>INV-M.1: INFORMATION EXTRACTION FROM SPEECH.....</b>	<b>3</b>
<i>John Makhoul, BBN Technologies, United States</i>	

<b>INV-M.2: GRAPH-BASED METHODS FOR LANGUAGE PROCESSING AND INFORMATION RETRIEVAL .....</b>	<b>4</b>
<i>Dragomir R. Radev, University of Michigan, United States</i>	

<b>INV-M.3: VOICE-ACTIVATED QUESTION ANSWERING .....</b>	<b>5</b>
<i>Dan Moldovan, Language Computer Corporation, United States</i>	

<b>INV-M.4: SPOKEN LANGUAGE UNDERSTANDING FOR CONVERSATIONAL DIALOG SYSTEMS .....</b>	<b>6</b>
<i>Michael McTear, University of Ulster, United Kingdom</i>	

## **INV-T: TUESDAY INVITED TUTORIALS**

<b>INV-T.1: APPLICATIONS OF SPOKEN LANGUAGE TECHNOLOGY AND SYSTEMS .....</b>	<b>7</b>
<i>Mike Philips, Mobius Corporation, United States</i>	

<b>INV-T.2: DESIGNING VOICE USER INTERFACES: EVIDENCE FROM THE FIELD .....</b>	
<i>Elizabeth Strand, Tellme Networks, United States</i>	

<b>INV-T.3: UNDERSTANDING AND MODELING COMMUNICATION SCENES .....</b>	<b>14</b>
<i>Herve Bourlard, IDIAP Research Institute, Switzerland</i>	

<b>INV-T.4: MULTILINGUAL LANGUAGE PROCESSING .....</b>	
<i>Pascale Fung, Hong Kong University of Science and Technology, Hong Kong SAR of China</i>	

## **INV-W: WEDNESDAY INVITED TUTORIALS**

### **INV-W.1: WIDENING THE NLP PIPELINE FOR SPOKEN LANGUAGE PROCESSING .....15**

*Srinivas Bangalore, AT&T Labs - Research, United States*

### **INV-W.2: SPOKEN LANGUAGE GENERATION**

*Johanna Moore, University of Edinburgh, United Kingdom*

### **INV-W.3: RECENT ADVANCES IN AUTOMATIC SPEECH SUMMARIZATION .....16**

*Sadaaki Furui, Tokyo Institute of Technology, Japan*

## **MP-1: INFORMATION EXTRACTION AND RETREIVAL**

### **MP-1.1: A NOVEL DTW-BASED DISTANCE MEASURE FOR SPEAKER SEGMENTATION .....22**

*Alex Park, James Glass, MIT, United States*

### **MP-1.2: IMPROVED SUMMARIZATION OF CHINESE SPOKEN DOCUMENTS BY PROBABILISTIC LATENT SEMANTIC ANALYSIS (PLSA) WITH FURTHER ANALYSIS AND INTEGRATED SCORING .....26**

*Sheng-yi Kong, Lin-shan Lee, National Taiwan University, Taiwan*

### **MP-1.3: DOMAIN-INDEPENDENT TOPIC SEGMENTATION USING A STRING KERNEL ON RECOGNIZED SUB-WORD SEQUENCES .....30**

*Ken Sadohara, Shi-wook Lee, Hiroaki Kojima, National Institute of Advanced Industrial Science and Technology (AIST), Japan*

### **MP-1.4: SUMMARIZATION OF SPOKEN LECTURES BASED ON LINGUISTIC SURFACE AND PROSODIC INFORMATION .....34**

*Shingo Togashi, Masaru Yamaguchi, Seiichi Nakagawa, Toyohashi University of Technology, Japan*

### **MP-1.5: SIMULATION ANALYSIS FOR INTERACTIVE RETRIEVAL OF SPOKEN DOCUMENTS WITH KEY TERMS RANKED BY REINFORCEMENT LEARNING .....38**

*Yi-cheng Pan, Lin-shan Lee, National Taiwan University, Taiwan*

### **MP-1.6: DISCRIMINATIVELY TRAINED SPOKEN DOCUMENT SIMILARITY MODELS AND THEIR APPLICATION TO PROBABILISTIC LATENT SEMANTIC ANALYSIS .....42**

*Kishan Thambiratnam, Frank Seide, Peng Yu, Microsoft Research Asia, China*

### **MP-1.7: INTEGRATION OF METADATA IN SPOKEN DOCUMENT SEARCH USING POSITION SPECIFIC POSTERIOR LATICES .....46**

*Jorge Silva, University of Southern California, United States; Ciprian Chelba, Google, United States; Alex Acero, Microsoft Corporation, United States*

### **MP-1.8: JUST-IN-TIME LATENT SEMANTIC ADAPTATION ON LANGUAGE MODEL FOR CHINESE SPEECH RECOGNITION USING WEB DATA .....50**

*Qin Gao, XiaoJun Lin, Xihong Wu, Peking University, China*

<b>MP-1.9: THE (NON)UTILITY OF LINGUISTIC FEATURES FOR PREDICTING PROMINENCE IN SPONTANEOUS SPEECH</b>	<b>54</b>
<i>Jason M. Brenier, Ani Nenkova, Anubha Kothari, Laura Whitton, David Beaver, Dan Jurafsky, Stanford University, United States</i>	
<b>MP-1.10: IMPACT OF AUTOMATIC COMMA PREDICTION ON POS/NAME TAGGING OF SPEECH</b>	<b>58</b>
<i>Dustin Hillard, University of Washington, United States; Zhongqiang Huang, Purdue University, United States; Heng Ji, Ralph Grishman, New York University, United States; Dilek Hakkani-Tür, International Computer Science Institute (ICSI), United States; Mary Harper, Purdue University, United States; Mari Ostendorf, University of Washington, United States; Wen Wang, SRI International, United States</i>	
<b>MP-1.11: SPOKEN QA BASED ON A PASSAGE RETRIEVAL ENGINE</b>	<b>62</b>
<i>Emilio Sanchis, Davide Buscaldi, Sergio Grau, LLuis Hurtado, David Griol, Universidad Politécnic de Valencia, Spain</i>	
<b>MP-2: SPOKEN LANGUAGE UNDERSTANDING</b>	
<b>MP-2.1: JOINTLY PREDICTING DIALOG ACT AND NAMED ENTITY FOR SPOKEN LANGUAGE UNDERSTANDING</b>	<b>66</b>
<i>Minwoo Jeong, Gary Geunbae Lee, POSTECH, Republic of Korea</i>	
<b>MP-2.2: DIALOGUE-ACT TAGGING USING SMART FEATURE SELECTION; RESULTS ON MULTIPLE CORPORA</b>	<b>70</b>
<i>Daan Verbree, Rutger Rienks, Dirk Heylen, University of Twente, Netherlands</i>	
<b>MP-2.3: ACOUSTIC-SYNTACTIC MAXIMUM ENTROPY MODEL FOR AUTOMATIC PROSODY LABELING</b>	<b>74</b>
<i>Vivek Kumar Rangarajan Sridhar, University of Southern California, United States; Srinivas Bangalore, AT&amp;T Research Labs, United States; Shrikanth Narayanan, University of Southern California, United States</i>	
<b>MP-2.4: A DBN-BASED MULTI-LEVEL STOCHASTIC SPOKEN LANGUAGE UNDERSTANDING SYSTEM</b>	<b>78</b>
<i>Fabrice Lefevre, LIA - University of Avignon, France</i>	
<b>MP-2.5: SPOKEN OPINION EXTRACTION FOR DETECTING VARIATIONS IN USER SATISFACTION</b>	<b>82</b>
<i>Frederic Bechet, University of Avignon, France; Geraldine Damnati, France Telecom R&amp;D, France; Nathalie Camelin, Renato de Mori, University of Avignon, France</i>	
<b>MP-2.6: UNSUPERVISED COMBINATION OF METRICS FOR SEMANTIC CLASS INDUCTION</b>	<b>86</b>
<i>Elias Iosif, Athanasios Tegos, Apostolos Pangos, Technical University of Crete, Greece; Eric Fosler-Lussier, Ohio State University, United States; Alexandros Potamianos, Technical University of Crete, Greece</i>	
<b>MP-2.7: PARSE STRUCTURE AND SEGMENTATION FOR IMPROVING SPEECH RECOGNITION</b>	<b>90</b>
<i>William McNeill, Jeremy Kahn, Dustin Hillard, Mari Ostendorf, SSLI Lab, United States</i>	
<b>MP-2.8: MODEL ADAPTATION FOR DIALOG ACT TAGGING</b>	<b>94</b>
<i>Gokhan Tur, Umit Guz, SRI International, United States; Dilek Hakkani-Tür, International Computer Science Institute (ICSI), United States</i>	

<b>MP-2.9: AUTOMATIC TOPIC SEGMENTATION AND LABELING IN MULTIPARTY DIALOGUE</b>	<b>98</b>
<i>Pei-Yun Hsueh, Johanna D. Moore, University of Edinburgh, United Kingdom</i>	
<b>MP-2.10: MODEL ADAPTATION FOR SENTENCE SEGMENTATION FROM SPEECH</b>	<b>102</b>
<i>Sébastien Cuendet, Dilek Hakkani-Tür, International Computer Science Institute (ICSI), United States; Gokhan Tur, SRI International, United States</i>	
<b>MP-2.11: ON THE EFFECT OF WORD ERROR RATE ON AUTOMATIC QUALITY MONITORING</b>	<b>106</b>
<i>George Saon, Bhuvana Ramabhadran, Geoffrey Zweig, IBM T. J. Watson Research Center, United States</i>	
<b>MP-2.12: EXPLOITING WORD-LEVEL FEATURES FOR EMOTION PREDICTION</b>	<b>110</b>
<i>Greg Nicholas, Mihai Rotaru, Diane Litman, University of Pittsburgh, United States</i>	
 <b>TP-1: SPOKEN AND MULTIMODAL DIALOG SYSTEMS AND APPLICATIONS</b>	
<b>TP-1.1: MISSING DATA ASR WITH FUSION OF FEATURES AND COMBINATION OF RECOGNIZERS</b>	<b>114</b>
<i>Neil Joshi, Ling Guan, Ryerson University, Canada</i>	
<b>TP-1.2: COMMUNICATION WITH ROBOTS: EVIDENCE FROM A WEB-BASED EXPERIMENT ON HUMAN-COMPUTER INTERACTION</b>	<b>118</b>
<i>Petra Gieselmann, University of Karlsruhe, Germany; Prisca Stenneken, Cath. University Eichstätt-Ingolstadt, Germany</i>	
<b>TP-1.3: UNIFIED TRAINING OF WFSM THROUGH A GENERIC INTERFACE</b>	<b>122</b>
<i>Mikel Penagarikano, Luis Javier Rodriguez, German Bordel, University of the Basque Country, Spain</i>	
<b>TP-1.4: STRUCTURAL REPRESENTATION OF THE PRONUNCIATION AND ITS USE FOR CALL</b>	<b>126</b>
<i>Nobuaki Minematsu, Satoshi Asakawa, Keikichi Hirose, The University of Tokyo, Japan</i>	
<b>TP-1.5: CONTEXT-SENSITIVE LANGUAGE MODELING FOR LARGE SETS OF PROPER NOUNS IN MULTIMODAL DIALOGUE SYSTEMS</b>	<b>130</b>
<i>Alexander Gruenstein, Stephanie Seneff, Massachusetts Institute of Technology, United States</i>	
<b>TP-1.6: A MULTIMODAL ARCHITECTURE FOR HOME CONTROL BY DISABLED USERS</b>	<b>134</b>
<i>Guillermo Pérez, Gabriel Amores, Pilar Manchón, Universidad de Sevilla, Spain</i>	
<b>TP-1.7: MODELLING THE GLOBAL ACOUSTIC CORRELATES OF EXPRESSIVITY FOR CHINESE TEXT-TO-SPEECH SYNTHESIS</b>	<b>138</b>
<i>Hongwu Yang, Tsinghua University, China; Helen Meng, Zhiyong Wu, Chinese University of Hong Kong, China; Lianhong Cai, Tsinghua University, China</i>	
<b>TP-1.8: BLENDING SPEECH AND VISUAL INPUT IN MULTIMODAL DIALOGUE SYSTEMS</b>	<b>142</b>
<i>Manolis Perakakis, Michail Toutoudakis, Alexandros Potamianos, Technical University of Crete, Greece</i>	

<b>TP-1.9: DYNAMIC VOCABULARY ADAPTATION FOR A DAILY AND REAL-TIME BROADCAST NEWS TRANSCRIPTION SYSTEM</b>	<b>146</b>
<i>Ciro Martins, L2F/INESC-ID - IEETA/UA, Portugal; António Teixeira, IEETA/UA, Portugal; João Neto, L2F/INESC-ID, Portugal</i>	
<b>TP-1.10: SHAPING TO CONVERGENCE: EXPERIMENTS WITH SPEECH GRAFFITI</b>	<b>150</b>
<i>Stefanie Tomko, Roni Rosenfeld, Carnegie Mellon University, United States</i>	
<b>TP-1.11: AN ASYNCHRONOUS DBN FOR AUDIO-VISUAL SPEECH RECOGNITION</b>	<b>154</b>
<i>Kate Saenko, Karen Livescu, MIT, United States</i>	
<b>TP-1.12: AN ACOUSTIC DOPPLER-BASED FRONT END FOR HANDS FREE SPOKEN USER INTERFACES</b>	<b>158</b>
<i>Kaustubh Kalgaonkar, Georgia Institute of Technology, United States; Bhiksha Raj, Mitsubishi Electric Research Laboratories, United States</i>	
<b>TP-1.13: GRAPHICAL MODEL REPRESENTATIONS OF WORD LATTICES</b>	<b>162</b>
<i>Gang Ji, Jeff Bilmes, University of Washington, United States; Jeff Michels, Stanford University, United States; Katrin Kirchhoff, University of Washington, United States; Chris Manning, Stanford University, United States</i>	
<b>TP-1.14: AUDIO-VISUAL BIOMETRIC RECOGNITION BY VECTOR QUANTIZATION</b>	<b>166</b>
<i>Amitava Das, Prasanta Ghosh, Microsoft Research - India, India</i>	
 <b>TP-2: DIALOG MANAGEMENT</b>	
<b>TP-2.1: ONLINE SUPERVISED LEARNING OF NON-UNDERSTANDING RECOVERY POLICIES</b>	<b>170</b>
<i>Dan Bohus, Brian Langner, Antoine Raux, Alan Black, Maxine Eskenazi, Alex Rudnicky, Carnegie Mellon University, United States</i>	
<b>TP-2.2: DIALOGUE CONTEXT-BASED RE-RANKING OF ASR HYPOTHESES</b>	<b>174</b>
<i>Rebecca Jonson, GU Dialogue Systems Lab &amp; GSLT, Sweden</i>	
<b>TP-2.3: EVALUATING EFFECTIVENESS AND PORTABILITY OF REINFORCEMENT LEARNED DIALOGUE STRATEGIES WITH REAL USERS: THE TALK TOWNINFO EVALUATION</b>	<b>178</b>
<i>Oliver Lemon, Kallirroi Georgila, James Henderson, Edinburgh University, United Kingdom</i>	
<b>TP-2.4: REINFORCEMENT LEARNING OF DIALOGUE STRATEGIES WITH HIERARCHICAL ABSTRACT MACHINES</b>	<b>182</b>
<i>Heriberto Cuayáhuatl, Steve Renals, Oliver Lemon, Hiroshi Shimodaira, University of Edinburgh, United Kingdom</i>	
<b>TP-2.5: AN ANALYSIS OF AUTOMATIC CONTENT SELECTION ALGORITHMS FOR SPOKEN DIALOGUE SYSTEM SUMMARIES</b>	<b>186</b>
<i>Joseph Polifroni, Marilyn Walker, University of Sheffield, United Kingdom</i>	

<b>TP-2.6: USING LOGISTIC REGRESSION TO INITIALISE</b> .....	<b>190</b>
<b>REINFORCEMENT-LEARNING-BASED DIALOGUE SYSTEMS</b>	
<i>Verena Rieser, Saarland University, Germany; Oliver Lemon, Edinburgh University, United Kingdom</i>	
<b>TP-2.7: CHAT AND GOAL-ORIENTED DIALOG TOGETHER: A UNIFIED</b> .....	<b>194</b>
<b>EXAMPLE-BASED ARCHITECTURE FOR MULTI-DOMAIN DIALOG MANAGEMENT</b>	
<i>Cheongjae Lee, Sangkeun Jung, Minwoo Jeong, Gary Geunbae Lee, POSTECH, Republic of Korea</i>	
<b>TP-2.8: VALUE-BASED OPTIMAL DECISION FOR DIALOG SYSTEMS</b> .....	<b>198</b>
<i>Esther Levin, City College of New York, United States; Roberto Pieraccini, SpeechCycle, United States</i>	
<b>TP-2.9: A SPOKEN DIALOG SYSTEM WITH AUTOMATIC RECOVERY</b> .....	<b>202</b>
<b>MECHANISM FROM MISRECOGNITION</b>	
<i>Norihide Kitaoka, Hirotoshi Yano, Seiichi Nakagawa, Toyohashi University of Technology, Japan</i>	
<b>TP-2.10: INTERACTION IN TASK-ORIENTED HUMAN-HUMAN DIALOGUE:</b> .....	<b>206</b>
<b>THE EFFECTS OF DIFFERENT TURN-TAKING POLICIES</b>	
<i>Raquel Fernandez, Tatjana Lucht, Kepa Rodriguez, David Schlangen, University of Potsdam, Germany</i>	
<b>TP-2.11: DIALOG SYSTEMS FOR SURVEYS: THE RATE-A-COURSE</b> .....	<b>210</b>
<b>SYSTEM</b>	
<i>Amanda Stent, Svetlana Stenchikova, Matthew Marge, Stony Brook University, United States</i>	
<b>TP-2.12: PLANNING AND LOGIC PROGRAMMING FOR DIALOG</b> .....	<b>214</b>
<b>MANAGEMENT</b>	
<i>Paul Fodor, Stony Brook University, United States; Juan Huerta, IBM T. J. Watson Research Center, United States</i>	
<b>TP-2.13: LET'S DISCOH: COLLECTING AN ANNOTATED OPEN CORPUS</b> .....	<b>218</b>
<b>WITH DIALOGUE ACTS AND REWARD SIGNALS FOR NATURAL LANGUAGE HELPDESKS</b>	
<i>Giovanni Andreani, ICSI / Speech Village, Italy; Giuseppe Di Fabbrizio, Mazin Gilbert, AT&amp;T Labs, Inc. - Research, United States; Dan Gillick, Dilek Hakkani-Tür, International Computer Science Institute (ICSI), United States; Oliver Lemon, Edinburgh University, United Kingdom</i>	
 <b>WP-1: MACHINE TRANSLATION</b>	
<b>WP-1.1: TOWARDS SPEECH TRANSLATION OF NON WRITTEN</b> .....	<b>222</b>
<b>LANGUAGES</b>	
<i>Laurent Besacier, Bowen Zhou, Yuqing Gao, IBM, United States</i>	
<b>WP-1.2: FOLSOM: A FAST AND MEMORY-EFFICIENT PHRASE-BASED</b> .....	<b>226</b>
<b>APPROACH TO STATISTICAL MACHINE TRANSLATION</b>	
<i>Bowen Zhou, Stanley Chen, Yuqing Gao, IBM T. J. Watson Research Center, United States</i>	
<b>WP-1.3: AUTOMATIC LEARNING OF CHINESE ENGLISH SEMANTIC</b> .....	<b>230</b>
<b>STRUCTURE MAPPING</b>	
<i>Pascale Fung, Zhaojun Wu, Yongsheng Yang, Dekai Wu, Hong Kong University of Science and Technology, Hong Kong SAR of China</i>	

<b>WP-1.4: INVERSION TRANSDUCTION GRAMMAR COVERAGE OF ARABIC-ENGLISH WORD ALIGNMENT FOR TREE-STRUCTURED STATISTICAL MACHINE TRANSLATION</b>	<b>234</b>
<i>Dekai Wu, Marine Carpuat, Yihai Shen, Hong Kong University of Science and Technology, Hong Kong SAR of China</i>	
<b>WP-1.5: SYNTACTIC PHRASE-BASED STATISTICAL MACHINE TRANSLATION</b>	<b>238</b>
<i>Hany Hassan, Mary Hearne, Andy Way, Dublin City University, Ireland; Khalil Sima'an, University of Amsterdam, Netherlands</i>	
<b>WP-1.6: REORDERING EXPERIMENTS FOR N-GRAM-BASED SMT</b>	<b>242</b>
<i>Josep M. Crego, Jose B. Mariño, TALP Research Center, Spain</i>	
<b>WP-1.7: MACHINE TRANSLATION SYSTEM DEVELOPMENT BASED ON HUMAN LIKENESS</b>	<b>246</b>
<i>Patrik Lambert, Jesús Giménez, Marta R. Costa-jussà, TALP Research Center, Spain; Enrique Amigó, Universidad Nacional de Educación a Distancia, Spain; Rafael Banchs, Lluís Màrquez, José A. Rodríguez Fonollosa, TALP Research Center, Spain</i>	
<b>WP-1.8: EXPERIMENTAL FACILITY FOR MEASURING THE IMPACT OF ENVIRONMENTAL NOISE AND SPEAKER VARIATION ON SPEECH-TO-SPEECH TRANSLATION DEVICES</b>	<b>250</b>
<i>Douglas Jones, Arvind Jairam, Wade Shen, MIT Lincoln Laboratory, United States; John Tardelli, Paul Gatewood, ARCON Corporation, United States; Michael Emonts, Defense Language Institute, United States</i>	
<b>WP-1.9: DESIGN AND EVALUATION OF THE 2006 BBN ENGLISH/IRAQI TWO-WAY SPEECH TRANSLATION SYSTEM</b>	<b>254</b>
<i>David Stallard, Fred Choi, Kriste Krstovski, Prem Natarajan, Rohit Prasad, Shirin Saleem, Raid Suleiman, BBN Technologies, United States</i>	
<b>WP-1.10: INTEGRATING TEXT AND PHONETIC INFORMATION FOR ROBUST STATISTICAL SPEECH TRANSLATION</b>	<b>258</b>
<i>Liang Gu, Yonggang Deng, Wei Zhang, Yuqing Gao, IBM, United States</i>	