

**9th Annual Conference of the  
International Speech  
Communication Association 2008  
(INTERSPEECH 2008)**

**Brisbane, Australia  
22-26 September 2008**

**Volume 1 of 5**

**ISBN: 978-1-61567-378-0**

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2008) by the International Speech Communications Association  
All rights reserved.

Printed by Curran Associates, Inc. (2009)

For permission requests, please contact the International Speech Communications Association  
at the address below.

International Speech Communications Association  
c/o Emmanuelle Foxonet  
Lous Tourils  
F-66390 Baixas, France

Phone: 33 468 385 827  
Fax: 49 228 735 639

[secretariat@isca-speech.org](mailto:secretariat@isca-speech.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

---

---

## TueSe2.O1 : Keynote 1: Hiroya Fujisaki — ISCA Medallist

Great Hall, Time 11:00 - 12:00, Tuesday 23rd September 2008

Chair: Isabel Trancoso

PAGE 1  
Keynote 1

### **In Search of Models in Speech Communication Research**

*Hiroya Fujisaki, University of Tokyo, Japan*

---

## WedSe1.O1 : Keynote 2: Abeer Alwan

Great Hall, Time 08:30 - 09:30, Wednesday 24th September 2008

Chair: Anne Cutler

PAGE 11  
Keynote 2

### **Dealing with Limited and Noisy Data in ASR: A Hybrid Knowledge-Based and Statistical Approach**

*Abeer Alwan, University of California at Los Angeles, USA*

---

## ThuSe1.O1 : Keynote 3: Joaquin Gonzalez-Rodriguez

Great Hall, Time 08:30 - 09:30, Thursday 25th September 2008

Chair: Michael Wagner

PAGE 16  
Keynote 3

### **Forensic Automatic Speaker Recognition: Fiction or Science?**

*Joaquin Gonzalez-Rodriguez, Universidad Autónoma de Madrid, Spain*

---

## FriSe1.O1 : Keynote 4: Justine Cassell

Great Hall, Time 08:30 - 09:30, Friday 26th September 2008

Chair: Denis Burnham

PAGE 18  
Keynote 4

### **Modeling Rapport in Embodied Conversational Agents**

*Justine Cassell, Northwestern University, USA*

---

---

---

## TueSe3.O1 : Segmentation and Classification

Great Hall, Time 13:30 - 15:30, Tuesday 23rd September 2008

Chair: Helen M. Meng

PAGE 20  
TueSe3.O1-1  
13:30 - 13:50

### **Agglomerative Hierarchical Speaker Clustering Using Incremental Gaussian Mixture Cluster Modeling**

*Kyu J. Han, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 24  
TueSe3.O1-2  
13:50 - 14:10

### **Weighted Segmental K-Means Initialization for SOM-Based Speaker Clustering**

*Oshry Ben-Harush<sup>1</sup>, Itshak Lapidot<sup>2</sup>, Hugo Guterman<sup>1</sup>*

*<sup>1</sup>Ben-Gurion University of the Negev, Israel; <sup>2</sup>Sami Shamoon College of Engineering, Israel*

PAGE 28  
TueSe3.O1-3  
14:10 - 14:30

### **Learning Essential Speaker Sub-Space Using Hetero-Associative Neural Networks for Speaker Clustering**

*Shajith Ikbal, Karthik Visweswariah, IBM India Research Lab, India*

PAGE 32  
TueSe3.O1-4  
14:30 - 14:50

### **Two's a Crowd: Improving Speaker Diarization by Automatically Identifying and Excluding Overlapped Speech**

*Kofi Boakye, Oriol Vinyals, Gerald Friedland, ICSI, USA*

PAGE 36  
TueSe3.O1-5  
14:50 - 15:10

### **T-Test Distance and Clustering Criterion for Speaker Diarization**

*Trung Hieu Nguyen<sup>1</sup>, Eng Siong Chng<sup>2</sup>, Haizhou Li<sup>1</sup>*

*<sup>1</sup>Institute for Infocomm Research, Singapore; <sup>2</sup>Nanyang Technological University, Singapore*

PAGE 40  
TueSe3.O1-6  
15:10 - 15:30

### **Integration of TDOA Features in Information Bottleneck Framework for Fast Speaker Diarization**

*Deepu Vijayaseenan, Fabio Valente, Hervé Bourlard, IDIAP Research Institute, Switzerland*

---

---

## TueSe3.O2: Speech Coding

Plaza 1, Time 13:30 – 15:30, Tuesday 23rd September 2008

Chair: Julien Epps

- PAGE 44  
TueSe3.O2-1  
13:30 – 13:50
- Low Complexity Near-Optimal Unit-Selection Algorithm for Ultra Low Bit-Rate Speech Coding Based on N-Best Lattice and Viterbi Search**  
*V. Ramasubramanian, D. Harish, Siemens Corporate Technology India, India*
- PAGE 45  
TueSe3.O2-2  
13:50 – 14:10
- A New Fast Algebraic Fixed Codebook Search Algorithm in CELP Speech Coding**  
*Václav Eksler<sup>1</sup>, Redwan Salami<sup>2</sup>, Milan Jelinek<sup>1</sup>*  
<sup>1</sup>University of Sherbrooke, Canada; <sup>2</sup>VoiceAge Corporation, Canada
- PAGE 49  
TueSe3.O2-3  
14:10 – 14:30
- A Novel Transcoding Algorithm Between 3GPP AMR-NB (7.95kbit/s) and ITU-T G.729a (8kbit/s)**  
*Hao Xu, Changchun Bao, Beijing University of Technology, China*
- PAGE 53  
TueSe3.O2-4  
14:30 – 14:50
- Mel-Frequency Cepstral Coefficient-Based Bandwidth Extension of Narrowband Speech**  
*Amr H. Nour-Eldin, Peter Kabal, McGill University, Canada*
- PAGE 57  
TueSe3.O2-5  
14:50 – 15:10
- A PCM Coding Noise Reduction for ITU-T G.711.1**  
*Jean-Luc Garcia, Claude Marro, Balazs Kövesi, Orange Labs, France*
- PAGE 61  
TueSe3.O2-6  
15:10 – 15:30
- An Instrumental Measure for End-to-End Speech Transmission Quality Based on Perceptual Dimensions: Framework and Realization**  
*Marcel Wältermann<sup>1</sup>, Kirstin Scholz<sup>2</sup>, Sebastian Möller<sup>1</sup>, Lu Huo<sup>2</sup>, Alexander Raake<sup>1</sup>, Ulrich Heute<sup>2</sup>*  
<sup>1</sup>Technische Universität Berlin, Germany; <sup>2</sup>Christian-Albrechts-Universität zu Kiel, Germany

---

---

## TueSe3.O3: Human Conversation and Communication

Plaza 2, Time 13:30 – 15:30, Tuesday 23rd September 2008

Chair: Bernd Möbius

- PAGE 65  
TueSe3.O3-1  
13:30 – 13:50
- Duration and F0 Interval of Utterance-Final Intonation Contours in the Perception of German Sentence Modality**  
*Benno Peters, Hartmut R. Pfizinger, Christian-Albrechts-Universität zu Kiel, Germany*
- PAGE 69  
TueSe3.O3-2  
13:50 – 14:10
- Contrastive Utterances Make Alternatives Salient — Cross-Modal Priming Evidence**  
*Bettina Braun, Lara Tagliapietra, Anne Cutler, Max Planck Institute for Psycholinguistics, The Netherlands*
- PAGE 70  
TueSe3.O3-3  
14:10 – 14:30
- Exploring a Mechanism of Speech Synchronization Using Auditory Delayed Experiments**  
*Masato Ishizaki<sup>1</sup>, Yasuharu Den<sup>2</sup>, Senshi Fukashiro<sup>1</sup>*  
<sup>1</sup>University of Tokyo, Japan; <sup>2</sup>Chiba University, Japan
- PAGE 74  
TueSe3.O3-4  
14:30 – 14:50
- Prosodic Manifestations of Confidence and Uncertainty in Spoken Language**  
*Heather Pon-Barry, Harvard University, USA*
- PAGE 78  
TueSe3.O3-5  
14:50 – 15:10
- Identifying Relevant Phrases to Summarize Decisions in Spoken Meetings**  
*Raquel Fernández<sup>1</sup>, Matthew Frampton<sup>1</sup>, John Dowding<sup>2</sup>, Anish Adukuzhiyil<sup>1</sup>, Patrick Ehlen<sup>1</sup>, Stanley Peters<sup>1</sup>*  
<sup>1</sup>Stanford University, USA; <sup>2</sup>University of California at Santa Cruz, USA
- PAGE 82  
TueSe3.O3-6  
15:10 – 15:30
- Recovering Participant Identities in Meetings from a Probabilistic Description of Vocal Interaction**  
*Kornel Laskowski<sup>1</sup>, Tanja Schultz<sup>2</sup>*  
<sup>1</sup>Universität Karlsruhe (TH), Germany; <sup>2</sup>Carnegie Mellon University, USA

---

---

## TueSe3.O4: Special Session: OzPhon08 — Phonetics and Phonology of Australian Aboriginal Languages

Plaza 3&4, Time 13:30 – 15:35, Tuesday 23rd September 2008

Chair: Marija Tabain

---

PAGE 86  
TueSe3.O4-1  
13:30 – 13:55

### Coarticulation in Nasal and Lateral Clusters in Warlpiri

*Janet Fletcher*<sup>1</sup>, *Deborah Loakes*<sup>1</sup>, *Andrew Butcher*<sup>2</sup>

<sup>1</sup>University of Melbourne, Australia; <sup>2</sup>Flinders University, Australia

PAGE 90  
TueSe3.O4-2  
13:55 – 14:20

### Phonetically Prestopped Laterals in Australian Languages: A Preliminary Investigation of Warlpiri

*Deborah Loakes*<sup>1</sup>, *Andrew Butcher*<sup>2</sup>, *Janet Fletcher*<sup>1</sup>, *Hywel Stoakes*<sup>1</sup>

<sup>1</sup>University of Melbourne, Australia; <sup>2</sup>Flinders University, Australia

PAGE 94  
TueSe3.O4-3  
14:20 – 14:45

### Connected Speech Processes in Warlpiri

*John Ingram*, *Mary Laughren*, *Jeff Chapman*, University of Queensland, Australia

PAGE 95  
TueSe3.O4-4  
14:45 – 15:10

### Consonant Enhancement in Lamalama, an Initial-Dropping Language of Cape York Peninsula, North Queensland

*Christina Pentland*, University of Queensland, Australia

PAGE 96  
TueSe3.O4-5  
15:10 – 15:35

### Text, Rhythm and Metrical Form in an Aboriginal Song Series

*Myfany Turpin*, University of Queensland, Australia

---

---

---

## TueSe3.P1: Acoustic Activity Detection, Pitch Tracking and Analysis

Mezzanine Level Area A1, Time 13:30 – 15:30, Tuesday 23rd September 2008

Chair: Timothy J. Hazen

---

PAGE 99  
TueSe3.P1-1

### Statistical Speech Activity Detection Based on Spatial Power Distribution for Analyses of Poster Presentations

*Kentaro Ishizuka*<sup>1</sup>, *Shoko Araki*<sup>1</sup>, *Tatsuya Kawahara*<sup>2</sup>

<sup>1</sup>NTT Corporation, Japan; <sup>2</sup>Kyoto University, Japan

PAGE 103  
TueSe3.P1-2

### A Statistical Model-Based Voice Activity Detection Employing Minimum Classification Error Technique

*Sang-Ick Kang*, *Ji-Hyun Song*, *Kye-Hwan Lee*, *Yun-Sik Park*, *Joon-Hyuk Chang*, Inha University, Korea

PAGE 107  
TueSe3.P1-3

### Comparative Evaluation of Different Methods for Voice Activity Detection

*Hongfei Ding*, *Koichi Yamamoto*, *Masami Akamine*, Toshiba Corporate R&D Center, Japan

PAGE 111  
TueSe3.P1-4

### Speech/Non-Speech Segments Detection Based on Chaotic and Prosodic Features

*Soheil Shafiee*, *Farshad Almasganj*, *Ayyoob Jafari*, Amirkabir University of Technology, Iran

PAGE 115  
TueSe3.P1-5

### Acoustic Event Classification Using a Distributed Microphone Network with a GMM/SVM Combined Algorithm

*Christian Zieger*, *Maurizio Omologo*, FBK-irst, Italy

PAGE 119  
TueSe3.P1-6

### Intentional Voice Command Detection for Completely Hands-Free Speech Interface in Home Environments

*Yasunari Obuchi*, *Masahito Togami*, *Takashi Sumiyoshi*, Hitachi Ltd., Japan

*TueSe3.P1 continued ...*

PAGE 123  
TueSe3.P1-7

**Fusion of Audio and Video Modalities for Detection of Acoustic Events**  
*Taras Butko, Andrey Temko, Climent Nadeu, Cristian Canton, Universitat Politècnica de Catalunya, Spain*

PAGE 127  
TueSe3.P1-8

**DySANA: Dynamic Speech and Noise Adaptation for Voice Activity Detection**  
*Ron J. Weiss<sup>1</sup>, Trausti Kristjansson<sup>2</sup>*  
<sup>1</sup>Columbia University, USA; <sup>2</sup>Google Inc., USA

PAGE 131  
TueSe3.P1-9

**A Comprehensive Study on the Effects of Room Reverberation on Fundamental Frequency Estimation**  
*Rico Petrick<sup>1</sup>, Masashi Unoki<sup>2</sup>, Anish Mittal<sup>3</sup>, Carlos Segura<sup>4</sup>, Ruediger Hoffmann<sup>1</sup>*  
<sup>1</sup>Technische Universität Dresden, Germany; <sup>2</sup>JAIST, Japan; <sup>3</sup>IIT Roorkee, India; <sup>4</sup>Universitat Politècnica de Catalunya, Spain

PAGE 135  
TueSe3.P1-10

**A Hybrid Speech Signal Based Algorithm for Pitch Marking Using Finite State Machines**  
*H. Hussein, M. Wolff, O. Jokisch, F. Duckhorn, G. Strecha, Ruediger Hoffmann, Technische Universität Dresden, Germany*

PAGE 139  
TueSe3.P1-11

**Parameter Estimation Method of F0 Control Model for Singing Voices**  
*Yasunori Ohishi<sup>1</sup>, Hirokazu Kameoka<sup>2</sup>, Kunio Kashino<sup>2</sup>, Kazuya Takeda<sup>1</sup>*  
<sup>1</sup>Nagoya University, Japan; <sup>2</sup>NTT Corporation, Japan

PAGE 143  
TueSe3.P1-12

**An Algorithm for Multi-Pitch Tracking in Co-Channel Speech**  
*Srikanth Vishnubhotla, Carol Y. Espy-Wilson, University of Maryland, USA*

PAGE 147  
TueSe3.P1-13

**Multipitch Tracking Using a Factorial Hidden Markov Model**  
*Michael Wohlmayr, Franz Pernkopf, Graz University of Technology, Austria*

---

*TueSe3.P1 continued ...*

PAGE 151  
TueSe3.P1-14

**Cochannel Speech Separation Using Multi-Pitch Estimation and Model Based Voiced Sequential Grouping**  
*Ming Li, Chuan Cao, Di Wang, Ping Lu, Qiang Fu, Yonghong Yan, Chinese Academy of Sciences, China*

PAGE 155  
TueSe3.P1-15

**Crosscorrelation of Adjacent Spectra Enhances Fundamental Frequency Tracking**  
*Philippe Martin, UFR Linguistique, France*

## TueSe3.P2: Single- and Multichannel Speech Enhancement I

*Mezzanine Level Area A2, Time 13:30 – 15:30, Tuesday 23rd September 2008*

*Chair: Martin J. Russell*

PAGE 159  
TueSe3.P2-1

**Enhancement of Noisy Speech Recordings via Blind Source Separation**  
*Jiri Malek, Zbynek Koldovsky, Jindrich Zdansky, Jan Nouza, Technical University of Liberec, Czech Republic*

PAGE 163  
TueSe3.P2-2

**Studies on Estimation of the Number of Sources in Blind Source Separation**  
*Takaaki Ishibashi<sup>1</sup>, Hidetoshi Nakashima<sup>1</sup>, Hiromu Gotanda<sup>2</sup>*  
<sup>1</sup>Kumamoto National College of Technology, Japan; <sup>2</sup>Kinki University, Japan

PAGE 167  
TueSe3.P2-3

**Speech Enhancement Based on Hypothesized Wiener Filtering**  
*V. Ramasubramanian<sup>1</sup>, Deepak Vijaywargi<sup>2</sup>*  
<sup>1</sup>Siemens Corporate Technology India, India; <sup>2</sup>University of Washington, USA

PAGE 171  
TueSe3.P2-4

**Psychoacoustically-Motivated Adaptive  $\beta$ -Order Generalized Spectral Subtraction Based on Data-Driven Optimization**  
*Junfeng Li<sup>1</sup>, Hui Jiang<sup>2</sup>, Masato Akagi<sup>1</sup>*  
<sup>1</sup>JAIST, Japan; <sup>2</sup>York University, Canada

PAGE 175  
TueSe3.P2-5

**Two Stage Iterative Wiener Filtering for Speech Enhancement**  
*Krishna Nand K., T.V. Sreenivas, Indian Institute of Science, India*

PAGE 179  
TueSe3.P2-6

**Assessment of Correlation Between Objective Measures and Speech Recognition Performance in the Evaluation of Speech Enhancement**  
*Pei Ding, Jie Hao, Toshiba China R&D Center, China*

---

---

## TueSe3.P3 : Spoken Language Systems I

Mezzanine Level Area B3, Time 13:30 - 15:30, Tuesday 23rd September 2008

Chair: Sebastian Möller

---

PAGE 183  
TueSe3.P3-1

### **Predicting ASR Errors by Exploiting Barge-In Rate of Individual Users for Spoken Dialogue Systems**

*Kazunori Komatani, Tatsuya Kawahara, Hiroshi G. Okuno, Kyoto University, Japan*

PAGE 187  
TueSe3.P3-2

### **Expanding Vocabulary for Recognizing User's Abbreviations of Proper Nouns Without Increasing ASR Error Rates in Spoken Dialogue Systems**

*Masaki Katsumaru, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno, Kyoto University, Japan*

PAGE 191  
TueSe3.P3-3

### **Exploiting the ASR N-Best by Tracking Multiple Dialog State Hypotheses**

*Jason D. Williams, AT&T Labs Research, USA*

PAGE 195  
TueSe3.P3-4

### **A Spoken Language Interpretation Component for a Robot Dialogue System**

*Enes Makalic, Ingrid Zukerman, Michael Niemann, Monash University, Australia*

PAGE 199  
TueSe3.P3-5

### **MUESLI: Multiple Utterance Error Correction for a Spoken Language Interface**

*Federico Cesari, Horacio Franco, Gregory K. Myers, Harry Bratt, SRI International, USA*

PAGE 203  
TueSe3.P3-6

### **Methods to Optimize Transcription of On-Line Media**

*Sarah Conrod<sup>1</sup>, Sara Basson<sup>2</sup>, Dimitri Kanevsky<sup>2</sup>*

*<sup>1</sup>Cape Breton University, Canada; <sup>2</sup>IBM T.J. Watson Research Center, USA*

PAGE 207  
TueSe3.P3-7

### **Discrimination of Task-Related Words for Vocabulary Design of Spoken Dialog Systems**

*Akinori Ito<sup>1</sup>, Toyomi Meguro<sup>1</sup>, Shozo Makino<sup>1</sup>, Motoyuki Suzuki<sup>2</sup>*

*<sup>1</sup>Tohoku University, Japan; <sup>2</sup>University of Tokushima, Japan*

---

*TueSe3.P3 continued ...*

PAGE 211  
TueSe3.P3-8

**Dialog Management Using Weighted Finite-State Transducers**

*Chiori Hori<sup>1</sup>, Kiyonori Ohtake<sup>1</sup>, Teruhisa Misu<sup>1</sup>, Hideki Kashioka<sup>1</sup>, Satoshi Nakamura<sup>2</sup>*

*<sup>1</sup>NICT, Japan; <sup>2</sup>ATR-SLC, Japan*

PAGE 215  
TueSe3.P3-9

**Probabilistic Answer Selection Based on Conditional Random Fields for Spoken Dialog System**

*Yoshitaka Yoshimi, Ryota Kakitsuba, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, Nagoya Institute of Technology, Japan*

PAGE 219  
TueSe3.P3-10

**Let's Go Lab: A Platform for Evaluation of Spoken Dialog Systems with Real World Users**

*Maxine Eskenazi, Alan W. Black, Antoine Raux, Brian Langner, Carnegie Mellon University, USA*

PAGE 220  
TueSe3.P3-11

**The Impact of Language Dynamics on the Capitalization of Broadcast News**

*Fernando Batista<sup>1</sup>, Nuno Mamede<sup>2</sup>, Isabel Trancoso<sup>2</sup>*

*<sup>1</sup>INESC-ID/ISCTE, Portugal; <sup>2</sup>INESC-ID/IST, Portugal*

PAGE 224  
TueSe3.P3-12

**Lightly Supervised Acoustic Model Training on EPPS Recordings**

*Matthias Paulik, Alex Waibel, Universität Karlsruhe (TH), Germany*

PAGE 228  
TueSe3.P3-13

**Fast Call-Classification System Development Without In-Domain Training Data**

*Christophe Servan, Frederic Bechet, LIA, France*

PAGE 232  
TueSe3.P3-14

**iCNC and iROVER: The Limits of Improving System Combination with Classification?**

*Björn Hoffmeister, Ralf Schlüter, Hermann Ney, RWTH Aachen University, Germany*

PAGE 236  
TueSe3.P3-15

**System Combination for Spoken Language Understanding**

*Stefan Hahn, Patrick Lehnen, Hermann Ney, RWTH Aachen University, Germany*



---

---

## TueSe3.P4: Emotion and Expression I

Mezzanine Level Area B4, Time 13:30 - 15:30, Tuesday 23rd September 2008

Chair: Denis Burnham

---

PAGE 240  
TueSe3.P4-1

### **Multidimensional Features of Emotional Speech**

*Tomoko Suzuki*<sup>1</sup>, *Machiko Ikemoto*<sup>2</sup>, *Tomoko Sano*<sup>3</sup>, *Toshihiko Kinoshita*<sup>1</sup>

<sup>1</sup>Kansai Medical University, Japan; <sup>2</sup>Doshisha University, Japan; <sup>3</sup>Seibi Gakuen College, Japan

PAGE 241  
TueSe3.P4-2

### **Leveraging Emotion Detection Using Emotions from Yes-No Answers**

*Narjès Boufaden*, *Pierre Dumouchel*, CRIM, Canada

PAGE 245  
TueSe3.P4-3

### **Vowel Placement During Operatic Singing: 'Come si Parla' or 'Aggiustamento'?**

*Thomas J. Millhouse*, *Dianna T. Kenny*, University of Sydney, Australia

PAGE 249  
TueSe3.P4-4

### **Study on Strained Rough Voice as a Conveyer of Rage**

*Yumiko O. Kato*, *Yoshifumi Hirose*, *Takahiro Kamai*, Matsushita Electric Industrial Co. Ltd., Japan

PAGE 253  
TueSe3.P4-5

### **Integrating Rule and Template-Based Approaches for Emotional Malay Speech Synthesis**

*Mumtaz Begum*<sup>1</sup>, *Raja N. Aion*<sup>1</sup>, *Roziati Zainuddin*<sup>1</sup>, *Zuraidah M. Don*<sup>1</sup>, *Gerry Knowles*<sup>2</sup>

<sup>1</sup>University of Malaya, Malaysia; <sup>2</sup>Miquet Worldwide Sdn. Bhd., Malaysia

PAGE 257  
TueSe3.P4-6

### **The Expression and Perception of Emotions: Comparing Assessments of Self versus Others**

*Carlos Busso*, *Shrikanth S. Narayanan*, University of Southern California, USA

PAGE 261  
TueSe3.P4-7

### **On the Role of Acting Skills for the Collection of Simulated Emotional Speech**

*Emiel Kraemer*, *Marc Swerts*, Tilburg University, The Netherlands

*TueSe3.P4 continued ...*

**Detection of Security Related Affect and Behaviour in Passenger Transport**

*Björn Schuller<sup>1</sup>, Matthias Wimmer<sup>2</sup>, Dejan Arsic<sup>1</sup>, Tobias Moosmayr<sup>3</sup>,  
Gerhard Rigoll<sup>1</sup>*

*<sup>1</sup>Technische Universität München, Germany; <sup>2</sup>Waseda University, Japan; <sup>3</sup>BMW  
Group, Germany*

---

---

## TueSe4.01: Automatic Speech Recognition: Acoustic Models I

Great Hall, Time 16:00 - 18:00, Tuesday 23rd September 2008

Chair: Roger K. Moore

---

PAGE 269  
TueSe4.01-1  
16:00 - 16:20

### Soft Margin Estimation with Various Separation Levels for LVCSR

*Jinyu Li<sup>1</sup>, Zhi-Jie Yan<sup>2</sup>, Chin-Hui Lee<sup>1</sup>, Ren-Hua Wang<sup>2</sup>*

<sup>1</sup>Georgia Institute of Technology, USA; <sup>2</sup>University of Science & Technology of China, China

PAGE 273  
TueSe4.01-2  
16:20 - 16:40

### On the Equivalence of Gaussian and Log-Linear HMMs

*Georg Heigold, Patrick Lehnen, Ralf Schlüter, Hermann Ney, RWTH Aachen*

*University, Germany*

PAGE 277  
TueSe4.01-3  
16:40 - 17:00

### Generalization of Extended Baum-Welch Parameter Estimation for Discriminative Training and Decoding

*Dimitri Kanevsky<sup>1</sup>, Tara N. Sainath<sup>2</sup>, Bhuvana Ramabhadran<sup>1</sup>, David Nahamoo<sup>1</sup>*

<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>MIT, USA

PAGE 281  
TueSe4.01-4  
17:00 - 17:20

### An Ellipsoid Constrained Quadratic Programming Perspective to Discriminative Training of HMMs

*Peng Liu, Frank K. Soong, Microsoft Research Asia, China*

PAGE 285  
TueSe4.01-5  
17:20 - 17:40

### Discriminative Training of Variable-Parameter HMMs for Noise Robust Speech Recognition

*Dong Yu<sup>1</sup>, Li Deng<sup>1</sup>, Yifan Gong<sup>2</sup>, Alex Acero<sup>1</sup>*

<sup>1</sup>Microsoft Research, USA; <sup>2</sup>Microsoft Corporation, USA

PAGE 289  
TueSe4.01-6  
17:40 - 18:00

### Towards a Non-Parametric Acoustic Model: An Acoustic Decision Tree for Observation Probability Calculation

*Jasha Droppo<sup>1</sup>, Michael L. Seltzer<sup>1</sup>, Alex Acero<sup>1</sup>, Yu-Hsiang Bosco Chiu<sup>2</sup>*

<sup>1</sup>Microsoft Research, USA; <sup>2</sup>Carnegie Mellon University, USA

---

---

## TueSe4.02: Accent and Language Identification

Plaza 1, Time 16:00 - 18:00, Tuesday 23rd September 2008

Chair: Yuko Kinoshita

---

PAGE 293  
TueSe4.02-1  
16:00 - 16:20

### **Experiments with the ABI (Accents of the British Isles) Speech Corpus**

*Shona D'Arcy*<sup>1</sup>, *Martin J. Russell*<sup>2</sup>

<sup>1</sup>Trinity College Dublin, Ireland; <sup>2</sup>University of Birmingham, UK

PAGE 297  
TueSe4.02-2  
16:20 - 16:40

### **Politecnico di Torino System for the 2007 NIST Language Recognition Evaluation**

*Fabio Castaldo*<sup>1</sup>, *Emanuele Dalmasso*<sup>1</sup>, *Pietro Laface*<sup>1</sup>, *Daniele Colibro*<sup>2</sup>,  
*Claudio Vair*<sup>2</sup>

<sup>1</sup>Politecnico di Torino, Italy; <sup>2</sup>Loquendo, Italy

PAGE 301  
TueSe4.02-3  
16:40 - 17:00

### **Discriminative Training and Channel Compensation for Acoustic Language Recognition**

*Valiantsina Hubeika*, *Lukáš Burget*, *Pavel Matějka*, *Petr Schwarz*, *Brno University of Technology, Czech Republic*

PAGE 305  
TueSe4.02-4  
17:00 - 17:20

### **Comparison of Variable Selection Methods and Classifiers for Native Accent Identification**

*Tingyao Wu*, *Peter Karsmakers*, *Hugo Van hamme*, *Dirk Van Compernelle*, *Katholieke Universiteit Leuven, Belgium*

PAGE 309  
TueSe4.02-5  
17:20 - 17:40

### **A Comparison of Subspace Feature-Domain Methods for Language Recognition**

*W.M. Campbell*, *Douglas E. Sturim*, *Pedro A. Torres-Carrasquillo*, *Douglas A. Reynolds*,  
*MIT, USA*

PAGE 313  
TueSe4.02-6  
17:40 - 18:00

### **Context-Dependent Phone Models and Models Adaptation for Phonotactic Language Recognition**

*Mohamed Faouzi BenZeghiba*, *Jean-Luc Gauvain*, *Lori Lamel*, *LIMSI, France*

---

---

## TueSe4.03: Emotion and Expression II

Plaza 2, Time 16:00 - 18:00, Tuesday 23rd September 2008

Chair: Ailbhe Ní Chasaide

---

PAGE 317  
TueSe4.03-1  
16:00 - 16:20

### **Emotions and Articulatory Precision**

*Martijn Goudbeek, Jean Philippe Goldman, Klaus R. Scherer, University of Geneva, Switzerland*

PAGE 318  
TueSe4.03-2  
16:20 - 16:40

### **Assessing Agreement of Observer- and Self-Annotations in Spontaneous Multimodal Emotion Data**

*Khiet P. Truong, Mark A. Neerincx, David A. van Leeuwen, TNO-D&V, The Netherlands*

PAGE 322  
TueSe4.03-3  
16:40 - 17:00

### **Emotion Recognition in Spontaneous Emotional Speech for Anonymity-Protected Voice Chat Systems**

*Yoshiko Arimoto<sup>1</sup>, Hiromi Kawatsu<sup>2</sup>, Sumio Ohno<sup>1</sup>, Hitoshi Iida<sup>1</sup>*

*<sup>1</sup>Tokyo University of Technology, Japan; <sup>2</sup>NIJLA, Japan*

PAGE 326  
TueSe4.03-4  
17:00 - 17:20

### **Assigning Suitable Phrasal Tones and Pitch Accents by Sensing Affective Information from Text to Synthesize Human-Like Speech**

*Mostafa Al Masum Shaikh, Md. Khademul Islam Molla, Keikichi Hirose, University of Tokyo, Japan*

PAGE 330  
TueSe4.03-5  
17:20 - 17:40

### **Cross-Language Study of Vocal Correlates of Affective States**

*Irena Yanushevskaya, Ailbhe Ní Chasaide, Christer Gobl, Trinity College Dublin, Ireland*

PAGE 334  
TueSe4.03-6  
17:40 - 18:00

### **Gender-Related Differences in the Production and Perception of Emotion**

*Marc Swerts, Emiel Krahmer, Tilburg University, The Netherlands*

---

---

**TueSe4.O4: Special Session: PANZE 2008 — Phonetics and Phonology of Australian and New Zealand English**

*Plaza 3&4, Time 16:00 – 18:05, Tuesday 23rd September 2008*

*Chair: Felicity Cox*

---

PAGE 338  
TueSe4.O4-1  
16:00 – 16:25

**The English Pronunciation of Successive Groups of Maori Speakers**

*Catherine I. Watson<sup>1</sup>, Margaret Maclagan<sup>2</sup>, Jeanette King<sup>2</sup>, Ray Harlow<sup>3</sup>*

<sup>1</sup>*University of Auckland, New Zealand;* <sup>2</sup>*University of Canterbury, New Zealand;*

<sup>3</sup>*University of Waikato, New Zealand*

PAGE 342  
TueSe4.O4-2  
16:25 – 16:50

**Reversal of Short Front Vowel Raising in Australian English**

*Felicity Cox, Sallyanne Palethorpe, Macquarie University, Australia*

PAGE 346  
TueSe4.O4-3  
16:50 – 17:15

**GOOSE on the Move: A Study of /u/-Fronting in Australian News Speech**

*Jennifer Price, Monash University, Australia*

PAGE 347  
TueSe4.O4-4  
17:15 – 17:40

**The Vowels of Australian Aboriginal English**

*Andrew Butcher<sup>1</sup>, Victoria Anderson<sup>2</sup>*

<sup>1</sup>*Flinders University, Australia;* <sup>2</sup>*University of Hawai'i at Mānoa, USA*

PAGE 351  
TueSe4.O4-5  
17:40 – 18:05

**Perception and Production of /i:/, /ɪə/ and /e:/ in Australian English**

*Robert H. Mannell, Macquarie University, Australia*

---

---

## TueSe4.P1 : Speaker Recognition and Diarisation

Mezzanine Level Area A1, Time 16:00 - 18:00, Tuesday 23rd September 2008

Chair: Robbie Vogt

---

- PAGE 355  
TueSe4.P1-1      **An Expert System in Speaker Verification Task**  
*Zbyněk Zajíc, Lukáš Machlica, Aleš Padrta, Jan Vaněk, Vlasta Radová, University of West Bohemia, Czech Republic*
- PAGE 359  
TueSe4.P1-2      **Cascading Appearance-Based Features for Visual Speaker Verification**  
*David Dean, Sridha Sridharan, Patrick Lucey, Queensland University of Technology, Australia*
- PAGE 363  
TueSe4.P1-3      **Improved Novelty Detection for Online GMM Based Speaker Diarization**  
*Konstantin Markov, Satoshi Nakamura, ATR-SLC, Japan*
- PAGE 367  
TueSe4.P1-4      **Analysis of Impostor Tests with High Scores in NIST-SRE Context**  
*Salah Eddine Mezaache, Jean-François Bonastre, Driss Matrouf, LIA, France*
- PAGE 371  
TueSe4.P1-5      **Reinforced Temporal Structure Information for Embedded Utterance-Based Speaker Recognition**  
*Anthony Larcher<sup>1</sup>, Jean-François Bonastre<sup>1</sup>, John S.D. Mason<sup>2</sup>*  
*<sup>1</sup>LIA, France; <sup>2</sup>Swansea University, UK*
- PAGE 375  
TueSe4.P1-6      **Fast Search for Common Segments in Speech Signals for Speaker Verification**  
*Michael Gerber, Beat Pfister, ETH Zürich, Switzerland*
- PAGE 379  
TueSe4.P1-7      **Audio-Visual Multilevel Fusion for Speech and Speaker Recognition**  
*Girija Chetty, Michael Wagner, University of Canberra, Australia*
- PAGE 383  
TueSe4.P1-8      **Clustering Initialization Based on Spatial Information for Speaker Diarization of Meetings**  
*J. Luque, Carlos Segura, Javier Hernando, Universitat Politècnica de Catalunya, Spain*

---

---

## TueSe4.P2 : Single- and Multichannel Speech Enhancement II

Mezzanine Level Area A2, Time 16:00 - 18:00, Tuesday 23rd September 2008

Chair: John H.L. Hansen

---

PAGE 387  
TueSe4.P2-1

### **Effect of Compressing the Dynamic Range of the Power Spectrum in Modulation Filtering Based Speech Enhancement**

*James G. Lyons, Kuldip K. Paliwal, Griffith University, Australia*

PAGE 391  
TueSe4.P2-2

### **A Long State Vector Kalman Filter for Speech Enhancement**

*Stephen So, Kuldip K. Paliwal, Griffith University, Australia*

PAGE 395  
TueSe4.P2-3

### **Subspace Based Speech Enhancement Using Gaussian Mixture Model**

*Achintya Kundu, Saikat Chatterjee, T.V. Sreenivas, Indian Institute of Science, India*

PAGE 399  
TueSe4.P2-4

### **Generalized Parametric Spectral Subtraction Using Weighted Euclidean Distortion**

*Amit Das, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 403  
TueSe4.P2-5

### **Sudden Noise Reduction Based on GMM with Noise Power Estimation**

*Nobuyuki Miyake, Tetsuya Takiguchi, Yasuo Ariki, Kobe University, Japan*

PAGE 407  
TueSe4.P2-6

### **Speech Enhancement Using a Wiener Denoising Technique and Musical Noise Reduction**

*Md. Jahangir Alam<sup>1</sup>, Sid-Ahmed Selouani<sup>2</sup>, Douglas O'Shaughnessy<sup>1</sup>, Sofia Ben Jebara<sup>3</sup>*

*<sup>1</sup>Université du Québec, Canada; <sup>2</sup>Université du Moncton, Canada; <sup>3</sup>Ecole Supérieure des Communications de Tunis, Tunisia*

PAGE 411  
TueSe4.P2-7

### **Regularized Non-Negative Matrix Factorization with Temporal Dependencies for Speech Denoising**

*Kevin W. Wilson<sup>1</sup>, Bhiksha Raj<sup>1</sup>, Paris Smaragdis<sup>2</sup>*

*<sup>1</sup>Mitsubishi Electric Research Labs, USA; <sup>2</sup>Adobe Systems, USA*



---

*TueSe4.P2 continued ...*

PAGE 415  
TueSe4.P2-8

**ICA-Based MAP Speech Enhancement with Multiple Variable Speech Distribution Models**

*Xin Zou, Peter Jančovič, Münevver Köküer, Martin J. Russell, University of Birmingham, UK*

PAGE 419  
TueSe4.P2-9

**Source Separation Based on Binaural Cues and Source Model Constraints**

*Ron J. Weiss, Michael I. Mandel, Daniel P.W. Ellis, Columbia University, USA*

PAGE 423  
TueSe4.P2-10

**Maximum Kurtosis Beamforming with the Generalized Sidelobe Canceller**

*Kenichi Kumatani<sup>1</sup>, John McDonough<sup>2</sup>, Barbara Rauch<sup>2</sup>, Philip N. Garner<sup>1</sup>, Weifeng Li<sup>1</sup>, John Dines<sup>1</sup>*

*<sup>1</sup>IDIAP Research Institute, Switzerland; <sup>2</sup>Saarland University, Germany*

PAGE 427  
TueSe4.P2-11

**Noise Robust Speech Dereverberation Using Constrained Inverse Filter**

*Ken'ichi Furuya<sup>1</sup>, Akitoshi Kataoka<sup>2</sup>, Yoichi Haneda<sup>1</sup>*

*<sup>1</sup>NTT Corporation, Japan; <sup>2</sup>Ryukoku University, Japan*

PAGE 431  
TueSe4.P2-12

**A Dual Microphone Coherence Based Method for Speech Enhancement in Headsets**

*Mohsen Rahmani, Ahmad Akbari, Beghdad Ayad, Iran University of Science & Technology, Iran*

PAGE 435  
TueSe4.P2-13

**Sound Capture System and Spatial Filter for Small Devices**

*Ivan Tashev<sup>1</sup>, Slavy Mihov<sup>2</sup>, Tyler Gleghorn<sup>3</sup>, Alex Acero<sup>1</sup>*

*<sup>1</sup>Microsoft Research, USA; <sup>2</sup>Technical University of Sofia, Bulgaria; <sup>3</sup>Microsoft Corporation, USA*

PAGE 439  
TueSe4.P2-14

**An Effective Microphone Array Post-Filter in Arbitrary Environments**

*Ning Cheng, Wen-ju Liu, Peng Li, Bo Xu, Chinese Academy of Sciences, China*

---

*TueSe4.P2 continued ...*

PAGE 443  
TueSe4.P2-15

**Localization of Multiple Sound Sources Based on Inter-Channel Correlation Using a Distributed Microphone System**

*Kook Cho, Hajime Okumura, Takanobu Nishiura, Yoichi Yamashita, Ritsumeikan University, Japan*

PAGE 447  
TueSe4.P2-16

**A Frequency Domain Approach for Speech Enhancement with Directionality Using Compact Microphone Array**

*Heng Zhang, Qiang Fu, Yonghong Yan, Chinese Academy of Sciences, China*

---

---

## TueSe4.P3 : Spoken Language Systems II

Mezzanine Level Area B3, Time 16:00 - 18:00, Tuesday 23rd September 2008

Chair: Jerome Bellegarda

---

PAGE 451  
TueSe4.P3-1

**Question and Answer Database Optimization Using Speech Recognition Results**  
*Shota Takeuchi, Tobias Cincarek, Hiromichi Kawanami, Hiroshi Saruwatari, Kiyohiro Shikano, NAIST, Japan*

PAGE 455  
TueSe4.P3-2

**Development and Evaluation of Hands-Free Spoken Dialogue System for Railway Station Guidance**  
*Hiroshi Saruwatari, Yu Takahashi, Hiroyuki Sakai, Shota Takeuchi, Tobias Cincarek, Hiromichi Kawanami, Kiyohiro Shikano, NAIST, Japan*

PAGE 459  
TueSe4.P3-3

**Statistical Shared Plan-Based Dialog Management**  
*Amanda J. Stent, Srinivas Bangalore, AT&T Labs Research, USA*

PAGE 463  
TueSe4.P3-4

**When Calls Go Wrong: How to Detect Problematic Calls Based on Log-Files and Emotions?**  
*Ota Herm<sup>1</sup>, Alexander Schmitt<sup>2</sup>, Jackson Liscombe<sup>3</sup>*  
<sup>1</sup>Czech Technical University in Prague, Czech Republic; <sup>2</sup>University of Ulm, Germany; <sup>3</sup>SpeechCycle Inc., USA

PAGE 467  
TueSe4.P3-5

**Unsupervised Learning of Edit Parameters for Matching Name Variants**  
*Dan Gillick<sup>1</sup>, Dilek Hakkani-Tür<sup>2</sup>, Michael Levit<sup>2</sup>*  
<sup>1</sup>University of California at Berkeley, USA; <sup>2</sup>ICSI, USA

PAGE 471  
TueSe4.P3-6

**Detection of Repetitions in Spontaneous Speech in Dialogue Sessions**  
*Mert Cevik<sup>1</sup>, Fuliang Weng<sup>2</sup>, Chin-Hui Lee<sup>1</sup>*  
<sup>1</sup>Georgia Institute of Technology, USA; <sup>2</sup>Robert Bosch Corp., USA

---

*TueSe4.P3 continued ...*

PAGE 475  
TueSe4.P3-7

**Automatic Customer Feedback Processing: Alarm Detection in Open Question Spoken Messages**

*Nathalie Camelin<sup>1</sup>, Geraldine Damnati<sup>2</sup>, Frederic Bechet<sup>1</sup>, Renato De Mori<sup>1</sup>*

*<sup>1</sup>LIA, France; <sup>2</sup>Orange Labs, France*

PAGE 479  
TueSe4.P3-8

**Minimal Training Based Semantic Categorization in a Voice Activated Question Answering (VAQA) System**

*Mithun Balakrishna, Marta Tatu, Dan Moldovan, Lymba Corporation, USA*

PAGE 483  
TueSe4.P3-9

**User Study of the Bayesian Update of Dialogue State Approach to Dialogue Management**

*B. Thomson, M. Gašić, S. Keizer, F. Mairesse, J. Schatzmann, K. Yu, Steve Young, University of Cambridge, UK*

PAGE 487  
TueSe4.P3-10

**Extensibility Verification of Robust Domain Selection Against Out-of-Grammar Utterances in Multi-Domain Spoken Dialogue System**

*Satoshi Ikeda, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno, Kyoto University, Japan*

PAGE 491  
TueSe4.P3-11

**Improving Large Scale Alphanumeric String Recognition Using Redundant Information**

*Ea-Ee Jan<sup>1</sup>, Osamuyimen Stewart<sup>1</sup>, Raymond Co<sup>2</sup>, David Lubensky<sup>1</sup>*

*<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>IBM Canada Global Business Services, Canada*

PAGE 495  
TueSe4.P3-12

**SPRAAK: An Open Source “Speech Recognition and Automatic Annotation Kit”**

*Kris Demuyne, Jan Roelens, Dirk Van Compernelle, Patrick Wambacq, Katholieke Universiteit Leuven, Belgium*

---

*TueSe4.P3 continued ...*

PAGE 496  
TueSe4.P3-13

**Preliminary Evaluation of Speech/Sound Recognition for Telemedicine  
Application in a Real Environment**

*Michel Vacher<sup>1</sup>, Anthony Fleury<sup>2</sup>, Jean-François Serignat<sup>1</sup>, Norbert Noury<sup>2</sup>,  
Hubert Glasson<sup>1</sup>*

*<sup>1</sup>LIG, France; <sup>2</sup>TIMC-IMAG, France*

PAGE 500  
TueSe4.P3-14

**MobiDic — A Mobile Dictation and Notetaking Application**

*Markku Turunen, Aleksi Melto, Anssi Kainulainen, Jaakko Hakulinen, University of  
Tampere, Finland*

PAGE 504  
TueSe4.P3-15

**Automatic Speech Recognition for Scientific Purposes — webASR**

*Thomas Hain, Asmaa El Hannani, Stuart N. Wrigley, Vincent Wan, University of  
Sheffield, UK*

PAGE 508  
TueSe4.P3-16

**Evaluation of a Live Broadcast News Subtitling System for Portugues**

*Hugo Meinedo, Márcio Viveiros, João Neto, INESC-ID/IST, Portugal*

---

---

## TueSe4.P4 : Perception, Production, Discourse and Dialog

Mezzanine Level Area B4, Time 16:00 – 18:00, Tuesday 23rd September 2008

Chair: Deborah Loakes

---

PAGE 512  
TueSe4.P4-1

### **Recognizing and Modelling Regional Varieties of Swedish**

*Jonas Beskow<sup>1</sup>, Gösta Bruce<sup>2</sup>, Laura Enflo<sup>1</sup>, Björn Granström<sup>1</sup>, Susanne Schötz<sup>2</sup>*

<sup>1</sup>KTH, Sweden; <sup>2</sup>Lund University, Sweden

PAGE 516  
TueSe4.P4-2

### **Vowel Duration, Compression and Lengthening in Stressed Syllables in Central and Southern Varieties of Standard Italian**

*John Hajek, Mary Stevens, University of Melbourne, Australia*

PAGE 520  
TueSe4.P4-3

### **Acoustic Cues for the Perception of Intonation in Cantonese**

*Joan K.-Y. Ma<sup>1</sup>, Valter Ciocca<sup>2</sup>, Tara L. Whitehill<sup>1</sup>*

<sup>1</sup>University of Hong Kong, China; <sup>2</sup>University of British Columbia, Canada

PAGE 524  
TueSe4.P4-4

### **Perception of Dialectal Prosody**

*Adrian Leemann, Beat Siebenhaar, University of Berne, Switzerland*

PAGE 528  
TueSe4.P4-5

### **Does the McGurk Effect Rely on Processing Time Constraints?**

*Christian Kroos, Ashlie Dreves, University of Western Sydney, Australia*

PAGE 529  
TueSe4.P4-6

### **Exploring the Uncanny Valley Effect with Talking Heads**

*Takaaki Kuratate, Kathryn Ayers, Jeesun Kim, Denis Burnham, University of Western Sydney, Australia*

PAGE 530  
TueSe4.P4-7

### **How Do the Elderly Talk to a Natural Language Call Routing System?**

*Knut Kvale, Ragnhild Halvorsrud, Telenor Research & Innovation, Norway*

PAGE 534  
TueSe4.P4-8

### **Analysis of Relationship Between Impression of Human-to-Human Conversations and Prosodic Change and Its Modeling**

*Ryota Nishimura<sup>1</sup>, Norihide Kitaoka<sup>2</sup>, Seiichi Nakagawa<sup>1</sup>*

<sup>1</sup>Toyohashi University of Technology, Japan; <sup>2</sup>Nagoya University, Japan

---

*TueSe4.P4 continued ...*

PAGE 538  
TueSe4.P4-9

**Utterance-Level Normalization for Relative Articulation Rate Analysis**

*Tuomo Saarni, Jussi Hakokari, Jouni Isoaho, Tapio Salakoski, University of Turku, Finland*

PAGE 542  
TueSe4.P4-10

**Syntactic Complexity Induces Explicit Grounding in the MapTask Corpus**

*Martin Tietze, Vera Demberg, Johanna D. Moore, University of Edinburgh, UK*

PAGE 543  
TueSe4.P4-11

**Do Discourse Cues Facilitate Recall in Information Presentation Messages?**

*Andi Winterboer, Johanna D. Moore, Fernanda Ferreira, University of Edinburgh, UK*

PAGE 544  
TueSe4.P4-12

**Structured Heterogeneity of English Stress Variants**

*Noriko Hattori, Mie University, Japan*

PAGE 545  
TueSe4.P4-13

**A Method for Automatically Estimating F0 Model Parameters and a Speech Re-Synthesis Tool Using F0 Model and STRAIGHT**

*Shota Sato<sup>1</sup>, Taro Kimura<sup>2</sup>, Yasuo Horiuchi<sup>1</sup>, Masafumi Nishida<sup>1</sup>, Shingo Kuroiwa<sup>1</sup>, Akira Ichikawa<sup>1</sup>*

*<sup>1</sup>Chiba University, Japan; <sup>2</sup>Nintendo Co. Ltd., Japan*

---

---

## WedSe2.O1 : Single-Channel Speech Enhancement

Great Hall, Time 10:00 - 12:00, Wednesday 24th September 2008

Chair: Frank K. Soong

---

PAGE 549  
WedSe2.O1-1  
10:00 - 10:20

### Noise Driven Short-Time Phase Spectrum Compensation Procedure for Speech Enhancement

*Anthony P. Stark, Kamil K. Wójcicki, James G. Lyons, Kuldip K. Paliwal, Griffith University, Australia*

PAGE 553  
WedSe2.O1-2  
10:20 - 10:40

### A Phase-Averaged Model for the Relationship Between Noisy Speech, Clean Speech and Noise in the Log-Mel Domain

*Friedrich Faubel, John McDonough, Dietrich Klakow, Saarland University, Germany*

PAGE 557  
WedSe2.O1-3  
10:40 - 11:00

### Time and Frequency Dependent Amplification for Speech Intelligibility Enhancement in Noisy Environments

*Henk Brouckxon<sup>1</sup>, Werner Verhelst<sup>1</sup>, Bart De Schuymer<sup>2</sup>*  
*<sup>1</sup>Vrije Universiteit Brussel, Belgium; <sup>2</sup>Televic nv, Belgium*

PAGE 561  
WedSe2.O1-4  
11:00 - 11:20

### A Wavelet Based Speech Enhancement Method Using Noise Classification and Shaping

*Mahdi Mohammadi<sup>1</sup>, Behzad Zamani<sup>1</sup>, Babak NaserSharif<sup>2</sup>, Mohsen Rahmani<sup>1</sup>, Ahmad Akbari<sup>1</sup>*  
*<sup>1</sup>Iran University of Science & Technology, Iran; <sup>2</sup>University of Guilan, Iran*

PAGE 565  
WedSe2.O1-5  
11:20 - 11:40

### Speech Enhancement Based on Novel Two-Step *a priori* SNR Estimators

*Md. Jahangir Alam<sup>1</sup>, Douglas O'Shaughnessy<sup>1</sup>, Sid-Ahmed Selouani<sup>2</sup>*  
*<sup>1</sup>Université du Québec, Canada; <sup>2</sup>Université du Moncton, Canada*

PAGE 569  
WedSe2.O1-6  
11:40 - 12:00

### A Speech Enhancement Approach Using Piecewise Linear Approximation of an Explicit Model of Environmental Distortions

*Jun Du, Qiang Huo, Microsoft Research Asia, China*



---

---

## WedSe2.O2: Speech Synthesis Methods I

Plaza 1, Time 10:00 - 12:00, Wednesday 24th September 2008

Chair: Rolf Carlson

---

PAGE 573  
WedSe2.O2-1  
10:00 - 10:20

### **Articulatory Control of HMM-Based Parametric Speech Synthesis Driven by Phonetic Knowledge**

*Zhen-Hua Ling*<sup>1</sup>, *Korin Richmond*<sup>2</sup>, *Junichi Yamagishi*<sup>2</sup>, *Ren-Hua Wang*<sup>1</sup>

<sup>1</sup>University of Science & Technology of China, China; <sup>2</sup>University of Edinburgh, UK

PAGE 577  
WedSe2.O2-2  
10:20 - 10:40

### **Minimum Generation Error Training with Direct Log Spectral Distortion on LSPs for HMM-Based Speech Synthesis**

*Yi-Jian Wu*, *Keiichi Tokuda*, Nagoya Institute of Technology, Japan

PAGE 581  
WedSe2.O2-3  
10:40 - 11:00

### **Robustness of HMM-Based Speech Synthesis**

*Junichi Yamagishi*<sup>1</sup>, *Zhen-Hua Ling*<sup>2</sup>, *Simon King*<sup>1</sup>

<sup>1</sup>University of Edinburgh, UK; <sup>2</sup>University of Science & Technology of China, China

PAGE 585  
WedSe2.O2-4  
11:00 - 11:20

### **Improving Preselection in Unit Selection Synthesis**

*Alistair Conkie*, *Ann Syrdal*, *Yeon-Jun Kim*, *Mark Beutnagel*, AT&T Labs Research, USA

PAGE 589  
WedSe2.O2-5  
11:20 - 11:40

### **Efficient Join Cost Computation for Unit Selection Based TTS Systems**

*Feng Ding*<sup>1</sup>, *Jani Nurminen*<sup>2</sup>, *Jilei Tian*<sup>3</sup>

<sup>1</sup>Nokia Research Center, China; <sup>2</sup>Nokia Devices R&D, Finland; <sup>3</sup>Nokia Research Center, Finland

PAGE 593  
WedSe2.O2-6  
11:40 - 12:00

### **A Phonetic Assessment of Cross-Language Voice Conversion**

*Kayoko Yanagisawa*, *Mark Huckvale*, University College London, UK

---

---

## WedSe2.03: Speaking Style and Emotion Recognition

Plaza 2, Time 10:00 - 12:00, Wednesday 24th September 2008

Chair: Javier Hernando

---

PAGE 597  
WedSe2.03-1  
10:00 - 10:20

### **Abandoning Emotion Classes — Towards Continuous Emotion Recognition with Modelling of Long-Range Dependencies**

*Martin Wöllmer<sup>1</sup>, Florian Eyben<sup>1</sup>, Stephan Reiter<sup>2</sup>, Björn Schuller<sup>1</sup>, Cate Cox<sup>3</sup>, Ellen Douglas-Cowie<sup>3</sup>, Roddy Cowie<sup>3</sup>*

<sup>1</sup>Technische Universität München, Germany; <sup>2</sup>EB Automotive GmbH, Germany;

<sup>3</sup>Queen's University Belfast, UK

PAGE 601  
WedSe2.03-2  
10:20 - 10:40

### **Patterns, Prototypes, Performance: Classifying Emotional User States**

*Dino Seppi<sup>1</sup>, Anton Batliner<sup>2</sup>, Björn Schuller<sup>3</sup>, Stefan Steidl<sup>2</sup>, Thurid Vogt<sup>4</sup>, Johannes Wagner<sup>4</sup>, Laurence Devillers<sup>5</sup>, Laurence Vidrascu<sup>5</sup>, Noam Amir<sup>6</sup>, Vered Aharonson<sup>7</sup>*

<sup>1</sup>FBK-irst, Italy; <sup>2</sup>Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany;

<sup>3</sup>Technische Universität München, Germany; <sup>4</sup>University of Augsburg, Germany;

<sup>5</sup>LIMSI, France; <sup>6</sup>Tel Aviv University, Israel; <sup>7</sup>Tel Aviv Academic College of

Engineering, Israel

PAGE 605  
WedSe2.03-3  
10:40 - 11:00

### **Recognition of Stress in Speech Using Wavelet Analysis and Teager Energy Operator**

*Ling He<sup>1</sup>, Margaret Lech<sup>1</sup>, Sheeraz Memon<sup>1</sup>, Nicholas Allen<sup>2</sup>*

<sup>1</sup>RMIT University, Australia; <sup>2</sup>University of Melbourne, Australia

PAGE 609  
WedSe2.03-4  
11:00 - 11:20

### **Effects of Vocal Effort and Speaking Style on Text-Independent Speaker Verification**

*Elizabeth Shriberg<sup>1</sup>, Martin Graciarena<sup>1</sup>, Harry Bratt<sup>1</sup>, Andreas Kathol<sup>1</sup>, Sachin S. Kajarekar<sup>1</sup>, Huda Jameel<sup>1</sup>, Colleen Richey<sup>1</sup>, Fred Goodman<sup>2</sup>*

<sup>1</sup>SRI International, USA; <sup>2</sup>MITRE Corporation, USA

---

*WedSe2.03 continued ...*

PAGE 613  
WedSe2.03-5  
11:20 - 11:40

**Robustness of Prosodic Features to Voice Imitation**

*Mireia Farrús<sup>1</sup>, Michael Wagner<sup>1</sup>, Jan Anguita<sup>1</sup>, Javier Hernando<sup>2</sup>*

*<sup>1</sup>University of Canberra, Australia; <sup>2</sup>Universitat Politècnica de Catalunya, Spain*

PAGE 617  
WedSe2.03-6  
11:40 - 12:00

**Phonetic and Speaker Variations in Automatic Emotion Classification**

*Vidhyasaharan Sethu, Eliathamby Ambikairajah, Julien Epps, University of New South Wales, Australia*

---

---

**WedSe2.O4: Special Session: Cross-Linguistic and Developmental Issues in the Perception and Production of Lexical Tone**

*Plaza 3&4, Time 10:00 – 12:00, Wednesday 24th September 2008*

*Chair: Karen Mattock*

---

PAGE 621  
WedSe2.O4-1  
10:00 – 10:20

**Infants' Native and Nonnative Tone Perception**

*Karen Mattock, Lancaster University, UK*

PAGE 622  
WedSe2.O4-2  
10:20 – 10:40

**Language Experience Dependent Plasticity for Pitch Representation in the Human Brainstem**

*Ananthanarayan Krishnan, Jackson Gandour, Jayaganesh Swaminathan, Purdue University, USA*

PAGE 623  
WedSe2.O4-3  
10:40 – 11:00

**Development of Tone Perception and Tone Production in Cantonese-Learning Children Aged 2 to 5 Years**

*Valter Ciocca<sup>1</sup>, Vivian W.-K. Ip<sup>2</sup>*

*<sup>1</sup>University of British Columbia, Canada; <sup>2</sup>University of Hong Kong, China*

PAGE 624  
WedSe2.O4-4  
11:00 – 11:20

**Tone Hyperarticulation in Cantonese Infant-Directed Speech**

*Nan Xu, Denis Burnham, University of Western Sydney, Australia*

PAGE 625  
WedSe2.O4-5  
11:20 – 11:40

**Influences on Tone in Sepedi, a Southern Bantu Language**

*Sabine Zerbian<sup>1</sup>, Etienne Barnard<sup>2</sup>*

*<sup>1</sup>University of the Witwatersrand, South Africa; <sup>2</sup>CSIR, South Africa*

PAGE 626  
WedSe2.O4-6  
11:40 – 12:00

**An Acoustic-Phonetic Comparative Analysis of Osaka and Kagoshima Japanese Tonal Phenomena**

*Shunichi Ishihara, Australian National University, Australia*

---

---

## WedSe2.P1 : Special Session: Auditory-Inspired Spectro-Temporal Features I

Mezzanine Level Area A1, Time 10:00 - 12:00, Wednesday 24th September 2008

Chair: Tiago H. Falk

---

PAGE 630  
WedSe2.P1-1

**Modulation Spectrogram Features for Improved Speaker Diarization**  
*Oriol Vinyals, Gerald Friedland, ICSI, USA*

PAGE 634  
WedSe2.P1-2

**Spectro-Temporal Features for Robust Far-Field Speaker Identification**  
*Tiago H. Falk, Wai-Yip Chan, Queen's University, Canada*

PAGE 638  
WedSe2.P1-3

**Long-Term Spectro-Temporal Information for Improved Automatic Speech Emotion Classification**  
*Siqing Wu, Tiago H. Falk, Wai-Yip Chan, Queen's University, Canada*

PAGE 642  
WedSe2.P1-4

**A Comparative Study on AM and FM Features**  
*Yotaro Kubo<sup>1</sup>, Shigeki Okawa<sup>2</sup>, Akira Kurematsu<sup>1</sup>, Katsuhiko Shirai<sup>1</sup>*  
*<sup>1</sup>Waseda University, Japan; <sup>2</sup>Chiba Institute of Technology, Japan*

PAGE 646  
WedSe2.P1-5

**Dimensionality Reduction of Modulation Frequency Features for Speech Discrimination**  
*Maria Markaki, Yannis Stylianou, University of Crete, Greece*

PAGE 650  
WedSe2.P1-6

**Spectral Envelope Recovery Beyond the Nyquist Limit for High-Quality Manipulation of Speech Sounds**  
*Hideki Kawahara<sup>1</sup>, Masanori Morise<sup>2</sup>, Hideki Banno<sup>3</sup>, Toru Takahashi<sup>4</sup>, Ryuichi Nisimura<sup>1</sup>, Toshio Irino<sup>1</sup>*  
*<sup>1</sup>Wakayama University, Japan; <sup>2</sup>Kwansei Gakuin University, Japan; <sup>3</sup>Meijo University, Japan; <sup>4</sup>Kyoto University, Japan*

---

*WedSe2.P1 continued ...*

PAGE 654  
WedSe2.P1-7

**Adaptive-Order Fractional Fourier Transform Features for Speech Recognition**

*Hui Yin, Xiang Xie, Jingming Kuang, Beijing Institute of Technology, China*

PAGE 658  
WedSe2.P1-8

**Robust Front End Processing for Speech Recognition in Reverberant Environments: Utilization of Speech Characteristics**

*Rico Petrick<sup>1</sup>, Xugang Lu<sup>2</sup>, Masashi Unoki<sup>2</sup>, Masato Akagi<sup>2</sup>, Ruediger Hoffmann<sup>1</sup>*

*<sup>1</sup>Technische Universität Dresden, Germany; <sup>2</sup>JAIST, Japan*

---

---

## WedSe2.P2: Speech Coding, Quality Measurement and Auditory Modelling

Mezzanine Level Area A2, Time 10:00 - 12:00, Wednesday 24th September 2008

Chair: David B. Grayden

---

PAGE 662  
WedSe2.P2-1

### **High-Quality Analysis/Synthesis Method Based on Temporal Decomposition for Speech Modification**

*Binh Phu Nguyen, Takeshi Shibata, Masato Akagi, JAIST, Japan*

PAGE 666  
WedSe2.P2-2

### **Improved Frame Loss Recovery Using Closed-Loop Estimation of Very Low Bit Rate Side Information**

*Philippe Gournay, University of Sherbrooke, Canada*

PAGE 670  
WedSe2.P2-3

### **Predictability of STRFs in Auditory Cortex Neurons Depends on Stimulus Class**

*Max F.K. Happel<sup>1</sup>, Simon Müller<sup>1</sup>, Jörn Anemüller<sup>2</sup>, Frank W. Ohl<sup>1</sup>*

*<sup>1</sup>Leibniz Institute for Neurobiology, Germany; <sup>2</sup>Carl von Ossietzky Universität Oldenburg, Germany*

PAGE 671  
WedSe2.P2-4

### **Higher Layer Coding of Non-Speech Like Signals Using Factorial Pulse Codebook**

*Udar Mittal, James P. Ashley, Jonathan Gibbs, Motorola, USA*

PAGE 675  
WedSe2.P2-5

### **Spectral Noise Shaping: Improvements in Speech/Audio Codec Based on Linear Prediction in Spectral Domain**

*Sriram Ganapathy<sup>1</sup>, Petr Motlicek<sup>1</sup>, Hynek Hermansky<sup>1</sup>, Harinath Garudadri<sup>2</sup>*

*<sup>1</sup>IDIAP Research Institute, Switzerland; <sup>2</sup>Qualcomm Inc., USA*

PAGE 679  
WedSe2.P2-6

### **Introducing the Compression Wave Cochlear Amplifier**

*Matthew R. Flax, W. Harvey Holmes, University of New South Wales, Australia*

PAGE 683  
WedSe2.P2-7

### **Goldman-Hodgkin-Katz Cochlear Hair Cell Models — A Foundation for Nonlinear Cochlear Mechanics**

*Matthew R. Flax, W. Harvey Holmes, University of New South Wales, Australia*

---

*WedSe2.P2 continued ...*

PAGE 687  
WedSe2.P2-8

**A 8–32 kb/s Embedded Wideband Speech Coding Candidate for ITU-T EV-VBR Standardization**

*Changchun Bao, Hai-ting Li, Ze-xin Liu, Rui Fan, Heng Zhu, Mao-shen Jia, Rui Li, Beijing University of Technology, China*

PAGE 691  
WedSe2.P2-9

**A Novel Non-Intrusive Objective Speech Quality Measurement Based on GMM and SVR**

*Jing Wang, Juan Luo, Shenghui Zhao, Beijing Institute of Technology, China*

PAGE 695  
WedSe2.P2-10

**Decision Tree Based Frame Mode Selection for AMR-WB+**

*Jong Kyu Kim, Seung Seop Park, Chang Woo Han, Nam Soo Kim, Seoul National University, Korea*

PAGE 699  
WedSe2.P2-11

**Assessment of Objective Quality Measures for Speech Intelligibility**

*W.M. Liu<sup>1</sup>, K.A. Jellyman<sup>1</sup>, N.W.D. Evans<sup>2</sup>, John S.D. Mason<sup>1</sup>*

*<sup>1</sup>Swansea University, UK; <sup>2</sup>EURECOM, France*

PAGE 703  
WedSe2.P2-12

**Assessment of the Speech-Quality Dimension “Noisiness” for the Instrumental Estimation and Analysis of Telephone-Band Speech Quality**

*Kirstin Scholz<sup>1</sup>, Christine Kühnel<sup>2</sup>, Marcel Wältermann<sup>2</sup>, Sebastian Möller<sup>2</sup>, Ulrich Heute<sup>1</sup>*

*<sup>1</sup>Christian-Albrechts-Universität zu Kiel, Germany; <sup>2</sup>Technische Universität Berlin, Germany*

PAGE 707  
WedSe2.P2-13

**Intelligibility Evaluation of Ramsey-Derived Interleavers for Internet Voice Streaming with the iLBC Codec**

*Angel M. Gomez, José L. Carmona, Antonio M. Peinado, Victoria Sánchez, José A. Gonzalez, Universidad de Granada, Spain*



---

---

## WedSe2.P3: Accent and Language Recognition

Mezzanine Level Area B3, Time 10:00 - 12:00, Wednesday 24th September 2008

Chair: Dirk Van Compernelle

---

PAGE 711  
WedSe2.P3-1

**Language Identification on Code-Switching Utterances Using Multiple Cues**  
*Dau-Cheng Lyu, Ren-Yuan Lyu, Chang Gung University, Taiwan*

PAGE 715  
WedSe2.P3-2

**Target-Oriented Phone Selection from Universal Phone Set for Spoken Language Recognition**  
*Rong Tong<sup>1</sup>, Bin Ma<sup>1</sup>, Haizhou Li<sup>1</sup>, Eng Siong Chng<sup>2</sup>*  
*<sup>1</sup>Institute for Infocomm Research, Singapore; <sup>2</sup>Nanyang Technological University, Singapore*

PAGE 719  
WedSe2.P3-3

**The MITLL NIST LRE 2007 Language Recognition System**  
*Pedro A. Torres-Carrasquillo, Elliot Singer, W.M. Campbell, Terry Gleason, Alan McCree, Douglas A. Reynolds, Fred Richardson, Wade Shen, Douglas E. Sturim, MIT, USA*

PAGE 723  
WedSe2.P3-4

**Eigen-Channel Compensation and Discriminatively Trained Gaussian Mixture Models for Dialect and Accent Recognition**  
*Pedro A. Torres-Carrasquillo, Douglas E. Sturim, Douglas A. Reynolds, Alan McCree, MIT, USA*

PAGE 727  
WedSe2.P3-5

**Anchor-Model Fusion for Language Recognition**  
*Ignacio Lopez-Moreno, Daniel Ramos, Joaquin Gonzalez-Rodriguez, Doroteo T. Toledano, Universidad Autónoma de Madrid, Spain*

PAGE 731  
WedSe2.P3-6

**Introducing a FM Based Feature to Hierarchical Language Identification**  
*Bo Yin<sup>1</sup>, Tharmarajah Thiruvaran<sup>1</sup>, Eliathamby Ambikairajah<sup>1</sup>, Fang Chen<sup>2</sup>*  
*<sup>1</sup>University of New South Wales, Australia; <sup>2</sup>NICTA, Australia*

---

*WedSe2.P3 continued ...*

PAGE 735  
WedSe2.P3-7

**Dialect Classification via Discriminative Training**

*Yun Lei, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 739  
WedSe2.P3-8

**BUT Language Recognition System for NIST 2007 Evaluations**

*Pavel Matějka, Lukáš Burget, Ondřej Glembek, Petr Schwarz, Valiantsina Hubeika, Michal Fapoš, Tomáš Mikolov, Oldřich Plchot, Jan Černocký, Brno University of Technology, Czech Republic*

PAGE 743  
WedSe2.P3-9

**Advances in Phonotactic Language Recognition**

*Ondřej Glembek, Pavel Matějka, Lukáš Burget, Tomáš Mikolov, Brno University of Technology, Czech Republic*

PAGE 747  
WedSe2.P3-10

**Dialect Separation Assessment Using Log-Likelihood Score Distributions**

*Mahnoosh Mehrabani, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 751  
WedSe2.P3-11

**Study on Unique Pharyngeal and Uvular Consonants in Foreign Accented Arabic**

*Yousef A. Alotaibi, Khondaker Abdullah-Al-Mamun, Ghulam Muhammad, King Saud University, Saudi Arabia*

PAGE 755  
WedSe2.P3-12

**Automatic Accent Classification Using Ensemble Methods**

*Fukun Bi, Jian Yang, Dan Xu, University of Yunnan, China*

PAGE 759  
WedSe2.P3-13

**Foreign Accent Identification Based on Prosodic Parameters**

*Marina Piat, Dominique Fohr, Irina Illina, LORIA, France*

PAGE 763  
WedSe2.P3-14

**Dialect Recognition Using Adapted Phonetic Models**

*Wade Shen, Nancy Chen, Douglas A. Reynolds, MIT, USA*

PAGE 767  
WedSe2.P3-15

**Beyond Frame Independence: Parametric Modelling of Time Duration in Speaker and Language Recognition**

*Alan McCree, Fred Richardson, Elliot Singer, Douglas A. Reynolds, MIT, USA*

---

---

## WedSe2.P4: Prosody: Prosodic Structure, Paralinguistic, Non-linguistic and Other Cues

Mezzanine Level Area B4, Time 10:00 – 12:00, Wednesday 24th September 2008

Chair: Wolfgang Hess

---

PAGE 771  
WedSe2.P4-1

**Testing a Large Corpus of Natural Standard Arabic for Rhythm Class**  
*Liz Dockendorf, Dalal Almubayei, Matthew Benton, University of Texas at Arlington, USA*

PAGE 772  
WedSe2.P4-2

**A Comparison of Two Acoustic Measurement Approaches to the Rhythm Continuum of Natural Chinese and English Speech**  
*Matthew Benton, Liz Dockendorf, University of Texas at Arlington, USA*

PAGE 776  
WedSe2.P4-3

**A Study of Pitch Patterns of Japanese English Analyzed via Comparative Linguistic Features of English and Japanese**  
*Tomoko Nariai, Kazuyo Tanaka, University of Tsukuba, Japan*

PAGE 780  
WedSe2.P4-4

**A Corpus-Based Prosodic Study of Alsatian, Belgian and Swiss French**  
*Cécile Woehrling, Philippe Boula de Mareüil, Martine Adda-Decker, Lori Lamel, LIMSI, France*

PAGE 784  
WedSe2.P4-5

**Prosodic Position Effects and Function Words in English: A Pilot Study**  
*Mitsuhiro Nakamura, Nihon University, Japan*

PAGE 785  
WedSe2.P4-6

**How Useful Are Polynomials for Analyzing Intonation?**  
*Laura E. de Ruiter, Max Planck Institute for Psycholinguistics, The Netherlands*

PAGE 789  
WedSe2.P4-7

**Adaptive Filter Based Prosody Modification Approach**  
*Qingcai Chen, Shusen Zhou, Dandan Wang, Xiaohong Yang, Harbin Institute of Technology, China*

---

*WedSe2.P4 continued ...*

PAGE 793  
WedSe2.P4-8

**Speech/Laughter Classification in Meeting Audio**

*Swe Zin Kalayar Khine, Tin Lay Nwe, Haizhou Li, Institute for Infocomm Research, Singapore*

PAGE 797  
WedSe2.P4-9

**Getting the Last Laugh: Automatic Laughter Segmentation in Meetings**

*Mary Tai Knox, Nelson Morgan, Nikki Mirghafori, ICSI, USA*

PAGE 801  
WedSe2.P4-10

**The Influence of Audio Presentation Style on Multitasking During Teleconferences**

*Stuart N. Wrigley, Simon Tucker, Guy J. Brown, Steve Whittaker, University of Sheffield, UK*

PAGE 805  
WedSe2.P4-11

**Balancing Spoken Content Adaptation and Unit Length in the Recognition of Emotion and Interest**

*Bogdan Vlasenko<sup>1</sup>, Björn Schuller<sup>2</sup>, Kinfe Tadesse Mengistu<sup>1</sup>, Gerhard Rigoll<sup>2</sup>, Andreas Wendemuth<sup>1</sup>*

*<sup>1</sup>Otto-von-Guericke-Universität Magdeburg, Germany; <sup>2</sup>Technische Universität München, Germany*

PAGE 809  
WedSe2.P4-12

**Nonverbal Responses to Social Inclusion and Exclusion**

*Emiel Kraemer, Juliëtte Schaafsma, Marc Swerts, Ad Vingerhoets, Tilburg University, The Netherlands*

PAGE 813  
WedSe2.P4-13

**Acoustic Analysis of Imitated Voice Produced by a Professional Impersonator**

*Tatsuya Kitamura, Konan University, Japan*

PAGE 817  
WedSe2.P4-14

**Detection of Speech Under Physical Stress: Model Development, Sensor Selection, and Feature Fusion**

*Sanjay A. Patil, John H.L. Hansen, University of Texas at Dallas, USA*

---

---

## WedSe3.01: Automatic Speech Recognition: Language Models I

Great Hall, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: Keikichi Hirose

---

PAGE 821  
WedSe3.01-1  
13:30 - 13:50

### **Improving Japanese Language Models Using POS Information**

*Langzhou Chen*<sup>1</sup>, *Hisayoshi Nagae*<sup>2</sup>, *Matt Stuttle*<sup>1</sup>

<sup>1</sup>*Toshiba Research Europe Ltd., UK*; <sup>2</sup>*Toshiba Corporate R&D Center, Japan*

PAGE 825  
WedSe3.01-2  
13:50 - 14:10

### **Discriminative N-Gram Language Modeling for Turkish**

*Ebru Arısoy*<sup>1</sup>, *Brian Roark*<sup>2</sup>, *Izhak Shafran*<sup>2</sup>, *Murat Saraçlar*<sup>1</sup>

<sup>1</sup>*Boğaziçi University, Turkey*; <sup>2</sup>*Oregon Health & Science University, USA*

PAGE 829  
WedSe3.01-3  
14:10 - 14:30

### **Rich Morphology Based N-Gram Language Models for Arabic**

*Ahmad Emami*, *Imed Zitouni*, *Lidia Mangu*, *IBM T.J. Watson Research Center, USA*

PAGE 833  
WedSe3.01-4  
14:30 - 14:50

### **Unsupervised Language Model Adaptation Based on Topic and Role Information in Multiparty Meetings**

*Songfang Huang*, *Steve Renals*, *University of Edinburgh, UK*

PAGE 837  
WedSe3.01-5  
14:50 - 15:10

### **Context Dependent Language Model Adaptation**

*X. Liu*, *M.J.F. Gales*, *P.C. Woodland*, *University of Cambridge, UK*

PAGE 841  
WedSe3.01-6  
15:10 - 15:30

### **Iterative Language Model Estimation: Efficient Data Structure & Algorithms**

*Bo-June Hsu*, *James Glass*, *MIT, USA*

---

---

## WedSe3.O2: Speaker Identification and Verification

Plaza 1, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: Jean-François Bonastre

---

PAGE 845  
WedSe3.O2-1  
13:30 - 13:50

### **Phone-Based Cepstral Polynomial SVM System for Speaker Recognition**

*Sachin S. Kajarekar, SRI International, USA*

PAGE 849  
WedSe3.O2-2  
13:50 - 14:10

### **Using MAP Estimation of Feature Transformation for Speaker Recognition**

*Donglai Zhu, Bin Ma, Haizhou Li, Institute for Infocomm Research, Singapore*

PAGE 853  
WedSe3.O2-3  
14:10 - 14:30

### **Factor Analysis Subspace Estimation for Speaker Verification with Short Utterances**

*Robbie Vogt, Brendan Baker, Sridha Sridharan, Queensland University of Technology, Australia*

PAGE 857  
WedSe3.O2-4  
14:30 - 14:50

### **Combining Continuous Progressive Model Adaptation and Factor Analysis for Speaker Verification**

*Mitchell McLaren<sup>1</sup>, Driss Matrouf<sup>1</sup>, Robbie Vogt<sup>2</sup>, Jean-François Bonastre<sup>1</sup>*

*<sup>1</sup>LIA, France; <sup>2</sup>Queensland University of Technology, Australia*

PAGE 861  
WedSe3.O2-5  
14:50 - 15:10

### **Adaptive Decision Tree-Based Phone Cluster Models for Speaker Clustering**

*Chia-Hsin Hsieh, Chung-Hsien Wu, Han-Ping Shen, National Cheng Kung University, Taiwan*

PAGE 865  
WedSe3.O2-6  
15:10 - 15:30

### **Speaker Recognition in Two-Wire Test Sessions**

*Hagai Aronowitz<sup>1</sup>, Yosef A. Solewicz<sup>2</sup>*

*<sup>1</sup>IBM Haifa Research Lab, Israel; <sup>2</sup>Bar-Ilan University, Israel*

---

---

## WedSe3.O3 : Prosodic Structure and Processing

Plaza 2, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: John Ingram

---

PAGE 869  
WedSe3.O3-1  
13:30 - 13:50

### **The Effect of Position on the Realization of Second Occurrence Focus**

*Jason B. Bishop, University of California at Los Angeles, USA*

PAGE 873  
WedSe3.O3-2  
13:50 - 14:10

### **Effects of Intonational Phrase Boundaries on Pitch-Accented Syllables in American English**

*Yen-Liang Shue<sup>1</sup>, Stefanie Shattuck-Hufnagel<sup>2</sup>, Markus Iseli<sup>1</sup>, Sun-Ah Jun<sup>1</sup>, Nanette Veilleux<sup>3</sup>, Abeer Alwan<sup>1</sup>*

*<sup>1</sup>University of California at Los Angeles, USA; <sup>2</sup>MIT, USA; <sup>3</sup>Simmons College, USA*

PAGE 877  
WedSe3.O3-3  
14:10 - 14:30

### **Examining Pitch-Accent Variability from an Exemplar-Theoretic Perspective**

*Michael Walsh, Katrin Schweitzer, Bernd Möbius, Hinrich Schütze, University of Stuttgart, Germany*

PAGE 881  
WedSe3.O3-4  
14:30 - 14:50

### **Correlation of Utterance Length and Segmental Duration in Finnish Is Questionable**

*Jussi Hakokari, Tuomo Saarni, Jouni Isoaho, Tapio Salakoski, University of Turku, Finland*

PAGE 885  
WedSe3.O3-5  
14:50 - 15:10

### **Different Roles of Pitch and Duration in Distinguishing Word Stress in English**

*Jiahong Yuan, Stephen Isard, Mark Liberman, University of Pennsylvania, USA*

PAGE 886  
WedSe3.O3-6  
15:10 - 15:30

### **Cross-Dialect Irish Prosody: Linguistic Constraints on Fujisaki Modelling**

*Maria O'Reilly, Ailbhe Ní Chasaide, Christer Gobl, Trinity College Dublin, Ireland*

---

---

## WedSe3.O4: Special Session: Auditory-Inspired Spectro-Temporal Features II

Plaza 3&4, Time 13:30 - 15:35, Wednesday 24th September 2008

Chair: Martin Heckmann

---

PAGE 890  
WedSe3.O4-1  
13:30 - 13:55

### **Introducing Temporal Asymmetries in Feature Extraction for Automatic Speech Recognition**

*G.S.V.S. Sivaram, Hynek Hermansky, IDIAP Research Institute, Switzerland*

PAGE 894  
WedSe3.O4-2  
13:55 - 14:20

### **A Closer Look on Hierarchical Spectro-Temporal Features (HIST)**

*Martin Heckmann, Xavier Domont, Frank Joublin, Christian Goerick, Honda Research Institute Europe GmbH, Germany*

PAGE 898  
WedSe3.O4-3  
14:20 - 14:45

### **Multi-Stream Spectro-Temporal Features for Robust Speech Recognition**

*Sherry Y. Zhao, Nelson Morgan, ICSI, USA*

PAGE 902  
WedSe3.O4-4  
14:45 - 15:10

### **The Value of Auditory Offset Adaptation and Appropriate Acoustic Modeling**

*Huan Wang<sup>1</sup>, David Gelbart<sup>2</sup>, Hans-Günter Hirsch<sup>3</sup>, Werner Hemmert<sup>4</sup>*

*<sup>1</sup>Infineon Technologies AG, Germany; <sup>2</sup>ICSI, USA; <sup>3</sup>Hochschule Niederrhein, Germany;*

*<sup>4</sup>Technische Universität München, Germany*

PAGE 906  
WedSe3.O4-5  
15:10 - 15:35

### **Optimization and Evaluation of Gabor Feature Sets for ASR**

*Bernd T. Meyer, Birger Kollmeier, Carl von Ossietzky Universität Oldenburg, Germany*



---

---

## WedSe3.P1 : Automatic Speech Recognition: Acoustic Models II

Mezzanine Level Area A1, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: Jean-Luc Gauvain

---

- PAGE 910  
WedSe3.P1-1      **A Shrinkage Estimator for Speech Recognition with Full Covariance HMMs**  
*Peter Bell, Simon King, University of Edinburgh, UK*
- PAGE 914  
WedSe3.P1-2      **Covariance Updates for Discriminative Training by Constrained Line Search**  
*Peter Bell, Simon King, University of Edinburgh, UK*
- PAGE 915  
WedSe3.P1-3      **Min-Max Discriminative Training of Decoding Parameters Using Iterative Linear Programming**  
*Brian Mak, Tom Ko, Hong Kong University of Science & Technology, China*
- PAGE 919  
WedSe3.P1-4      **Discriminative Training for Complementariness in System Combination**  
*Daniel Willett<sup>1</sup>, Chuang He<sup>2</sup>*  
*<sup>1</sup>Nuance Communications, Germany; <sup>2</sup>Nuance Communications, USA*
- PAGE 920  
WedSe3.P1-5      **Penalty Function Maximization for Large Margin HMM Training**  
*George Saon, Daniel Povey, IBM T.J. Watson Research Center, USA*
- PAGE 924  
WedSe3.P1-6      **Implicit State-Tying for Support Vector Machines Based Speech Recognition**  
*Daniel Bolaños, Wayne Ward, University of Colorado at Boulder, USA*
- PAGE 928  
WedSe3.P1-7      **Using KL-Based Acoustic Models in a Large Vocabulary Recognition Task**  
*Guillermo Aradilla, Hervé Bouchard, Mathew Magimai Doss, IDIAP Research Institute, Switzerland*
- PAGE 932  
WedSe3.P1-8      **Acoustic Modeling Based on Model Structure Annealing for Speech Recognition**  
*Sayaka Shiota, Kei Hashimoto, Heiga Zen, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, Nagoya Institute of Technology, Japan*

---

*WedSe3.P1 continued ...*

PAGE 936  
WedSe3.P1-9

**Bayesian Context Clustering Using Cross Valid Prior Distribution for HMM-Based Speech Recognition**

*Kei Hashimoto, Heiga Zen, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, Nagoya Institute of Technology, Japan*

PAGE 940  
WedSe3.P1-10

**Speech Recognition Using Soft Decision Trees**

*Jitendra Ajmera, Masami Akamine, Toshiba Corporate R&D Center, Japan*

PAGE 944  
WedSe3.P1-11

**GPU-Accelerated Gaussian Clustering for fMPE Discriminative Training**

*Yu Shi, Frank Seide, Frank K. Soong, Microsoft Research Asia, China*

PAGE 948  
WedSe3.P1-12

**Discriminative Training Using the Trusted Expectation Maximization**

*Yasser Hifny, Yuqing Gao, IBM T.J. Watson Research Center, USA*

PAGE 952  
WedSe3.P1-13

**Maximum Mutual Information Estimation with Unlabeled Data for Phonetic Classification**

*Jui-Ting Huang, Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign, USA*

PAGE 956  
WedSe3.P1-14

**Maximum Accept and Reject (MARS) Training of HMM-GMM Speech Recognition Systems**

*Vivek Tyagi, IBM India Research Lab, India*

PAGE 960  
WedSe3.P1-15

**Nonlinear Mixture Autoregressive Hidden Markov Models for Speech Recognition**

*Sundar Srinivasan<sup>1</sup>, Tao Ma<sup>1</sup>, Daniel May<sup>1</sup>, Georgios Lazarou<sup>2</sup>, Joseph Picone<sup>1</sup>*

*<sup>1</sup>Mississippi State University, USA; <sup>2</sup>New York City Transit Authority, USA*

PAGE 964  
WedSe3.P1-16

**GPU Accelerated Acoustic Likelihood Computations**

*Patrick Cardinal, Pierre Dumouchel, Gilles Boulianne, Michel Comeau, CRIM, Canada*

---

---

## WedSe3.P2: Robust Automatic Speech Recognition I

Mezzanine Level Area A2, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: Kuldip K. Paliwal

---

PAGE 968  
WedSe3.P2-1

### **CENSREC-4: Development of Evaluation Framework for Distant-Talking Speech Recognition Under Reverberant Environments**

*Masato Nakayama*<sup>1</sup>, *Takanobu Nishiura*<sup>1</sup>, *Yuki Denda*<sup>1</sup>, *Norihide Kitaoka*<sup>2</sup>, *Kazumasa Yamamoto*<sup>3</sup>, *Takeshi Yamada*<sup>4</sup>, *Satoru Tsuge*<sup>5</sup>, *Chiyomi Miyajima*<sup>2</sup>, *Masakiyo Fujimoto*<sup>6</sup>, *Tetsuya Takiguchi*<sup>7</sup>, *Satoshi Tamura*<sup>8</sup>, *Tetsuji Ogawa*<sup>9</sup>, *Shigeki Matsuda*<sup>10</sup>, *Shingo Kuroiwa*<sup>11</sup>, *Kazuya Takeda*<sup>2</sup>, *Satoshi Nakamura*<sup>10</sup>

<sup>1</sup>Ritsumeikan University, Japan; <sup>2</sup>Nagoya University, Japan; <sup>3</sup>Toyohashi University of Technology, Japan; <sup>4</sup>University of Tsukuba, Japan; <sup>5</sup>University of Tokushima, Japan; <sup>6</sup>NTT Corporation, Japan; <sup>7</sup>Kobe University, Japan; <sup>8</sup>Gifu University, Japan; <sup>9</sup>Waseda University, Japan; <sup>10</sup>ATR-SLC, Japan; <sup>11</sup>Chiba University, Japan

PAGE 972  
WedSe3.P2-2

### **In-Car Speech Recognition Using Model-Based Wiener Filter and Multi-Condition Training**

*Masanori Tsujikawa*, *Takayuki Arakawa*, *Ryosuke Isotani*, NEC Corporation, Japan

PAGE 976  
WedSe3.P2-3

### **Adaptive Beamforming and Soft Missing Data Decoding for Robust Speech Recognition in Reverberant Environments**

*Marco Kühne*<sup>1</sup>, *Roberto Togneri*<sup>1</sup>, *Sven Nordholm*<sup>2</sup>

<sup>1</sup>University of Western Australia, Australia; <sup>2</sup>WATRI, Australia

PAGE 980  
WedSe3.P2-4

### **Spectral Subtraction in Likelihood-Maximizing Framework for Robust Speech Recognition**

*Bagher BabaAli*<sup>1</sup>, *Hossein Sameti*<sup>2</sup>, *Mehran Safayani*<sup>2</sup>

<sup>1</sup>Islamic Azad University at Dashtestan, Iran; <sup>2</sup>Sharif University of Technology, Iran

---

*WedSe3.P2 continued ...*

PAGE 984  
WedSe3.P2-5

**Front-End for Far-Field Speech Recognition Based on Frequency Domain Linear Prediction**

*Sriram Ganapathy, Samuel Thomas, Hynek Hermansky, IDIAP Research Institute, Switzerland*

PAGE 988  
WedSe3.P2-6

**Mask Estimation Incorporating Time-Frequency Trajectories for a CASA-Based ASR Front-End**

*Ji Hun Park, Jae Sam Yoon, Hong Kook Kim, GIST, Korea*

PAGE 992  
WedSe3.P2-7

**Soft Missing-Feature Mask Generation for Simultaneous Speech Recognition System in Robots**

*Toru Takahashi<sup>1</sup>, Shun'ichi Yamamoto<sup>1</sup>, Kazuhiro Nakadai<sup>2</sup>, Kazunori Komatani<sup>1</sup>, Tetsuya Ogata<sup>1</sup>, Hiroshi G. Okuno<sup>1</sup>*

*<sup>1</sup>Kyoto University, Japan; <sup>2</sup>Honda Research Institute Japan Co. Ltd., Japan*

PAGE 996  
WedSe3.P2-8

**A Posterior Approach for Microphone Array Based Speech Recognition**

*Dong Wang, Ivan Himawan, Joe Frankel, Simon King, University of Edinburgh, UK*

PAGE 1000  
WedSe3.P2-9

**Analysis of Physiologically-Motivated Signal Processing for Robust Speech Recognition**

*Yu-Hsiang Bosco Chiu, Richard M. Stern, Carnegie Mellon University, USA*

PAGE 1004  
WedSe3.P2-10

**Evaluation of Modulation Spectrum Equalization Techniques for Large Vocabulary Robust Speech Recognition**

*Liang-che Sun, Chang-wen Hsu, Lin-shan Lee, National Taiwan University, Taiwan*

PAGE 1008  
WedSe3.P2-11

**Confusion-Based Entropy-Weighted Decoding for Robust Speech Recognition**

*Yi Chen, Chia-yu Wan, Lin-shan Lee, National Taiwan University, Taiwan*

---

*WedSe3.P2 continued ...*

PAGE 1012  
WedSe3.P2-12

**Cepstral Domain Voice Activity Detection for Improved Noise Modeling in MMSE Feature Enhancement for ASR**

*Svein Gunnar Pettersen, Magne Hallstein Johnsen, NTNU, Norway*

PAGE 1016  
WedSe3.P2-13

**Unsupervised Re-Scoring of Observation Probability Based on Maximum Entropy Criterion by Using Confidence Measure with Telephone Speech**

*Carlos Molina, Nestor Becerra Yoma, Fernando Huenupan, Claudio Garreton, Universidad de Chile, Chile*

PAGE 1020  
WedSe3.P2-14

**Within-Class Feature Normalization for Robust Speech Recognition**

*Yuan-Fu Liao<sup>1</sup>, Chi-Hui Hsu<sup>1</sup>, Chi-Min Yang<sup>1</sup>, Jeng-Shien Lin<sup>2</sup>, Sen-Chia Chang<sup>2</sup>*

*<sup>1</sup>National Taipei University of Technology, Taiwan; <sup>2</sup>ITRI, Taiwan*

PAGE 1024  
WedSe3.P2-15

***a posteriori* SNR Weighted Energy Based Variable Frame Rate Analysis for Speech Recognition**

*Zheng-Hua Tan, Børge Lindberg, Aalborg University, Denmark*

PAGE 1028  
WedSe3.P2-16

**Silence Feature Normalization for Robust Speech Recognition in Additive Noise Environments**

*Chieh-cheng Wang, Chi-an Pan, Jieh-weih Hung, National Chi Nan University, Taiwan*

PAGE 1032  
WedSe3.P2-17

**Blind Dereverberation Based on CMN and Spectral Subtraction by Multi-Channel LMS Algorithm**

*L. Wang<sup>1</sup>, Seiichi Nakagawa<sup>2</sup>, Norihide Kitaoka<sup>3</sup>*

*<sup>1</sup>Shizuoka University, Japan; <sup>2</sup>Toyohashi University of Technology, Japan; <sup>3</sup>Nagoya University, Japan*

---

---

## WedSe3.P3: Speech Analysis and Processing, Voice Conversion and Modification

Mezzanine Level Area B3, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: Douglas O'Shaughnessy

---

PAGE 1036  
WedSe3.P3-1

### **Amplitude and Amplitude Variation of Emotional Speech**

*Hartmut R. Pfitzinger, Christian Kaernbach, Christian-Albrechts-Universität zu Kiel, Germany*

PAGE 1040  
WedSe3.P3-2

### **Babble Speech: Acoustic and Perceptual Variability**

*Nitish Krishnamurthy, Ayako Ikeno, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1044  
WedSe3.P3-3

### **On the Properties of a Time-Varying Quasi-Harmonic Model of Speech**

*Yannis Pantazis<sup>1</sup>, Olivier Rosec<sup>2</sup>, Yannis Stylianou<sup>1</sup>*

*<sup>1</sup>University of Crete, Greece; <sup>2</sup>Orange Labs, France*

PAGE 1048  
WedSe3.P3-4

### **Extraction and Tracking of Formant Response Jitter in the Cochlea for Objective Prediction of SB/SF DAM Attributes**

*Wenliang Lu, D. Sen, University of New South Wales, Australia*

PAGE 1052  
WedSe3.P3-5

### **Consonant Discrimination of Degraded Speech Using an Efferent-Inspired Closed-Loop Cochlear Model**

*David P. Messing<sup>1</sup>, Lorraine Delhorne<sup>1</sup>, Ed Bruckert<sup>2</sup>, Louis D. Braid<sup>1</sup>, Oded Ghitza<sup>2</sup>*

*<sup>1</sup>MIT, USA; <sup>2</sup>Sensimetrics Corporation, USA*

PAGE 1056  
WedSe3.P3-6

### **On the Development of Variable Length Teager Energy Operator (VTEO)**

*Vikrant Tomar, Hemant A. Patil, DA-IICT, India*

PAGE 1060  
WedSe3.P3-7

### **Metric Learning for Unsupervised Phoneme Segmentation**

*Yu Qiao, Nobuaki Minematsu, University of Tokyo, Japan*

---

*WedSe3.P3 continued ...*

PAGE 1064  
WedSe3.P3-8

**Combining Task-Dependent Information with Auditory Attention Cues for Prominence Detection in Speech**

*Ozlem Kalinli, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 1068  
WedSe3.P3-9

**Probabilistic Feature Mapping Based on Trajectory HMMs**

*Heiga Zen, Yoshihiko Nankaku, Keiichi Tokuda, Nagoya Institute of Technology, Japan*

PAGE 1072  
WedSe3.P3-10

**Simultaneous Conversion of Duration and Spectrum Based on Statistical Models Including Time-Sequence Matching**

*Kaori Yutani<sup>1</sup>, Yosuke Uto<sup>1</sup>, Yoshihiko Nankaku<sup>1</sup>, Tomoki Toda<sup>2</sup>, Keiichi Tokuda<sup>1</sup>*

*<sup>1</sup>Nagoya Institute of Technology, Japan; <sup>2</sup>NAIST, Japan*

PAGE 1076  
WedSe3.P3-11

**Low-Delay Voice Conversion Based on Maximum Likelihood Estimation of Spectral Parameter Trajectory**

*Takashi Muramatsu, Yamato Ohtani, Tomoki Toda, Hiroshi Saruwatari, Kiyohiro Shikano, NAIST, Japan*

PAGE 1080  
WedSe3.P3-12

**An Improved One-to-Many Eigenvoice Conversion System**

*Yamato Ohtani, Tomoki Toda, Hiroshi Saruwatari, Kiyohiro Shikano, NAIST, Japan*

PAGE 1084  
WedSe3.P3-13

**Study on Manipulation Method of Voice Quality Based on the Vocal Tract Area Function**

*Yoshinori Uchimura<sup>1</sup>, Hideki Banno<sup>1</sup>, Fumitada Itakura<sup>1</sup>, Hideki Kawahara<sup>2</sup>*

*<sup>1</sup>Meijo University, Japan; <sup>2</sup>Wakayama University, Japan*

PAGE 1088  
WedSe3.P3-14

**Incorporating Durational Modification in Voice Transformation**

*Arthur Toth, Alan W. Black, Carnegie Mellon University, USA*

PAGE 1092  
WedSe3.P3-15

**Non-Segmental Duration Feature Extraction for Prosodic Classification**

*Amy Dashiell, Brian Hutchinson, Anna Margolis, Mari Ostendorf, University of Washington, USA*

---

---

## WedSe3.P4: Special Session: Tonality in Production and Perception, Language in Australia and New Zealand

Mezzanine Level Area B4, Time 13:30 - 15:30, Wednesday 24th September 2008

Chair: Yoshinori Sagisaka

---

- PAGE 1096  
WedSe3.P4-1      **An ERP Study on Categorical Perception of Lexical Tones and Nonspeech Pitches**  
*Hong-Ying Zheng, William S.-Y. Wang, Chinese University of Hong Kong, China*
- PAGE 1097  
WedSe3.P4-2      **The Role of Japanese Pitch Accent in Spoken-Word Recognition: Evidence from Middle-Aged Accentless Dialect Listeners**  
*Takashi Otake<sup>1</sup>, Marii Higuchi<sup>2</sup>*  
<sup>1</sup>E-Listening Laboratory, Japan; <sup>2</sup>Dokkyo University, Japan
- PAGE 1101  
WedSe3.P4-3      **Mandarin Chinese Tone Nucleus Detection with Landmarks**  
*Siwei Wang, Gina-Anne Levow, University of Chicago, USA*
- PAGE 1105  
WedSe3.P4-4      **A Comparative Study on Dissyllabic Stress Patterns of Mandarin and Cantonese**  
*Weixiang Hu<sup>1</sup>, Jin Jian<sup>2</sup>, Aijun Li<sup>3</sup>, Xia Wang<sup>1</sup>*  
<sup>1</sup>Nokia Research Center, China; <sup>2</sup>Sun Yat-sen University, China; <sup>3</sup>Chinese Academy of Social Sciences, China
- PAGE 1109  
WedSe3.P4-5      **Three-Sectional-Staff Characterization of Cantonese Level Tones**  
*Rerrario Shui-Ching Ho, Yoshinori Sagisaka, Waseda University, Japan*
- PAGE 1113  
WedSe3.P4-6      **A Seven-Tone Dialect in Southern China with Falling-Rising-Falling Contour: A Linguistic Acoustic Analysis**  
*Xiaonong Zhu, Caicai Zhang, Hong Kong University of Science & Technology, China*
- PAGE 1116  
WedSe3.P4-7      **Pitch Target Analysis of Thai Tones Using Quantitative Target Approximation Model and Unsupervised Clustering**  
*Santitham Prom-on, KMUTT, Thailand*



---

*WedSe3.P4 continued ...*

PAGE 1120  
WedSe3.P4-8

**Do English Speakers Assimilate Mandarin Tones to English Prosodic Categories?**  
*Connie K. So, Catherine T. Best, University of Western Sydney, Australia*

PAGE 1121  
WedSe3.P4-9

**Evidence of a Near-Merger in Western Sydney Australian English Vowels**  
*Rikke L. Bundgaard-Nielsen, Catherine T. Best, Michael D. Tyler, Christian Kroos, University of Western Sydney, Australia*

PAGE 1122  
WedSe3.P4-10

**Central Vowels in Arrernte: Metrical Prominence and Pitch Accent**  
*Marija Tabain<sup>1</sup>, Kristine Rickard<sup>1</sup>, Gavan Breen<sup>2</sup>, Veronica Dobson<sup>2</sup>*  
*<sup>1</sup>La Trobe University, Australia; <sup>2</sup>Institute for Aboriginal Development, Australia*

PAGE 1123  
WedSe3.P4-11

**Pausing and Phrase Length in Two Australian Languages**  
*Bella Ross, University of Melbourne, Australia*

PAGE 1124  
WedSe3.P4-12

**Positional Effects on the Characterization of Ejectives in Waima'a**  
*Mary Stevens, John Hajek, University of Melbourne, Australia*

PAGE 1128  
WedSe3.P4-13

**A Niuean Variant of New Zealand English?**  
*Donna Starks, Laura Thompson, Catherine I. Watson, University of Auckland, New Zealand*

---

---

## WedSe4.O1: Automatic Speech Recognition: Tone Languages

Great Hall, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Björn Granström

---

PAGE 1129  
WedSe4.O1-1  
16:00 - 16:20

**Phonetic Confusion Analysis and Robust Phone Set Generation for Shanghai-Accented Mandarin Speech Recognition**  
*Guo-Hong Ding, Nokia Research Center, China*

PAGE 1133  
WedSe4.O1-2  
16:20 - 16:40

**Prosody for Mandarin Speech Recognition: A Comparative Study of Read and Spontaneous Speech**  
*Yu Ting Yeung<sup>1</sup>, Yao Qian<sup>2</sup>, Tan Lee<sup>1</sup>, Frank K. Soong<sup>2</sup>*  
<sup>1</sup>Chinese University of Hong Kong, China; <sup>2</sup>Microsoft Research Asia, China

PAGE 1137  
WedSe4.O1-3  
16:40 - 17:00

**Improved Large Vocabulary Mandarin Speech Recognition by Selectively Using Tone Information with a Two-Stage Prosodic Model**  
*Li-Wei Cheng, Lin-shan Lee, National Taiwan University, Taiwan*

PAGE 1141  
WedSe4.O1-4  
17:00 - 17:20

**Mandarin Connected Digits Recognition for Whispered Speech**  
*Tingting Ru, Xiang Xie, Hui Yin, Jingming Kuang, Beijing Institute of Technology, China*

PAGE 1145  
WedSe4.O1-5  
17:20 - 17:40

**Recognizing Named Entities in Spoken Chinese Dialogues with a Character-Level Maximum Entropy Tagger**  
*Changchun Bao, Weiqun Xu, Yonghong Yan, Chinese Academy of Sciences, China*

PAGE 1149  
WedSe4.O1-6  
17:40 - 18:00

**A Novel Approach in Continuous Speech Recognition for Vietnamese, an Isolating Tonal Language**  
*Hong Quang Nguyen<sup>1</sup>, Pascal Nocera<sup>1</sup>, Eric Castelli<sup>2</sup>, Van Loan Trinh<sup>2</sup>*  
<sup>1</sup>LIA, France; <sup>2</sup>MICA, Vietnam

---

---

## WedSe4.O2: Spoken Dialogue Systems

Plaza 1, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Isabel Trancoso

---

PAGE 1153  
WedSe4.O2-1  
16:00 - 16:20

**Evaluating Semantic-Level Confidence Scores with Multiple Hypotheses**  
*B. Thomson, K. Yu, M. Gašić, S. Keizer, F. Mairesse, J. Schatzmann, Steve Young,*  
*University of Cambridge, UK*

PAGE 1157  
WedSe4.O2-2  
16:20 - 16:40

**Structured Models for Joint Decoding of Repeated Utterances**  
*Geoffrey Zweig, Dan Bohus, Xiao Li, Patrick Nguyen, Microsoft Research, USA*

PAGE 1161  
WedSe4.O2-3  
16:40 - 17:00

**A Bayesian Approach to Semantic Composition for Spoken Language Interpretation**  
*Marie-Jean Meurs, Fabrice Lefèvre, Renato De Mori, LIA, France*

PAGE 1165  
WedSe4.O2-4  
17:00 - 17:20

**Accommodating Explicit User Expressions of Uncertainty in Voice Search or Something Like That**  
*Tim Paek, Yun-Cheng Ju, Microsoft Research, USA*

PAGE 1169  
WedSe4.O2-5  
17:20 - 17:40

**Effects of User Modeling on POMDP-Based Dialogue Systems**  
*Dongho Kim<sup>1</sup>, Hyeong Seop Sim<sup>1</sup>, Kee-Eung Kim<sup>1</sup>, Jin Hyung Kim<sup>1</sup>,*  
*Hyunjeong Kim<sup>2</sup>, Joo Won Sung<sup>2</sup>*  
<sup>1</sup>KAIST, Korea; <sup>2</sup>KT, Korea

PAGE 1173  
WedSe4.O2-6  
17:40 - 18:00

**The Best of Both Worlds: Unifying Conventional Dialog Systems and POMDPs**  
*Jason D. Williams, AT&T Labs Research, USA*

---

---

## WedSe4.O3 : Cross-Language and Language-Specific Phonetics

Plaza 2, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Catherine T. Best

---

PAGE 1177  
WedSe4.O3-1  
16:20 - 16:40

### **The Assimilation of L2 Australian English Vowels to L1 Japanese Vowel**

**Categories: Vocabulary Size Matters**

*Rikke L. Bundgaard-Nielsen, Catherine T. Best, Michael D. Tyler, University of Western Sydney, Australia*

PAGE 1178  
WedSe4.O3-2  
16:40 - 17:00

### **Vowel Epenthesis, Acoustics and Phonology Patterns in Moroccan Arabic**

*Azra N. Ali<sup>1</sup>, Mohamed Lahrouchi<sup>2</sup>, Michael Ingleby<sup>1</sup>*

<sup>1</sup>University of Huddersfield, UK; <sup>2</sup>SFL, France

PAGE 1182  
WedSe4.O3-3  
17:00 - 17:20

### **Estimation of Vocal Tract Area Function for Mandarin Vowel Sequences Using MRI**

*Gaowu Wang<sup>1</sup>, Jianwu Dang<sup>1</sup>, Jiangping Kong<sup>2</sup>*

<sup>1</sup>JAIST, Japan; <sup>2</sup>Peking University, China

PAGE 1186  
WedSe4.O3-4  
17:20 - 17:40

### **The Effect of First Language (L1) Dialects on the Identification of Vietnamese Word-Final Stops**

*Kimiko Tsukada<sup>1</sup>, Thu T.A. Nguyễn<sup>2</sup>*

<sup>1</sup>Macquarie University, Australia; <sup>2</sup>University of Queensland, Australia

PAGE 1190  
WedSe4.O3-5  
17:40 - 18:00

### **Perceptual Evidence of Modern Greek Voiced Stops as Phonological Categories**

*Mark Antoniou, Catherine T. Best, Michael D. Tyler, University of Western Sydney, Australia*

PAGE 1191  
WedSe4.O3-6  
17:40 - 18:00

### **The Effect of Auditory and Visual Degradation on Audiovisual Perception of Native and Non-Native Speakers**

*Valerie Hazan, Enid Li, University College London, UK*

---

---

## WedSe4.O4: Special Session: Prosody of Spontaneous Speech I

Plaza 3&4, Time 16:00 – 18:00, Wednesday 24th September 2008

Chair: Hartmut R. Pfitzinger

---

PAGE 1195  
WedSe4.O4-1

**Quantitative Prosodic Analysis of Spontaneous Speech**  
*Hansjörg Mixdorff, TFH Berlin, Germany*

PAGE 1196  
WedSe4.O4-2

**The Effect of Cognitive Load on Disfluencies During In-Vehicle Spoken Dialogue**  
*Anders Lindström<sup>1</sup>, Jessica Villing<sup>2</sup>, Staffan Larsson<sup>2</sup>, Alexander Seward<sup>3</sup>,  
Nina Åberg<sup>4</sup>, Cecilia Holtelius<sup>5</sup>*  
<sup>1</sup>TeliaSonera, Sweden; <sup>2</sup>University of Gothenburg, Sweden; <sup>3</sup>Veridict AB, Sweden;  
<sup>4</sup>Volvo Technology AB, Sweden; <sup>5</sup>Volvo Car Corporation AB, Sweden

PAGE 1200  
WedSe4.O4-3

**Discourse Prosody Context — Global F0 and Tempo Modulations**  
*Chiu-yu Tseng, Zhao-yu Su, Academia Sinica, Taiwan*

PAGE 1204  
WedSe4.O4-4

**A Method for Automatic and Dynamic Estimation of Discourse Genre Typology with Prosodic Features**  
*Nicolas Obin<sup>1</sup>, Anne Lacheret-Dujour<sup>2</sup>, Christophe Veaux<sup>1</sup>, Xavier Rodet<sup>1</sup>,  
Anne-Catherine Simon<sup>3</sup>*  
<sup>1</sup>IRCAM, France; <sup>2</sup>MoDyCo, France; <sup>3</sup>Université Catholique de Louvain, Belgium

PAGE 1208  
WedSe4.O4-5

**The Meanings Carried by Interjections in Spontaneous Speech**  
*Carlos T. Ishi, Hiroshi Ishiguro, Norihiro Hagita, ATR-IRC, Japan*

PAGE 1212  
WedSe4.O4-6

**Speech Interaction with an Emotional Robotic Dog**  
*Christian M. Jones<sup>1</sup>, Andrew Deeming<sup>2</sup>*  
<sup>1</sup>University of the Sunshine Coast, Australia; <sup>2</sup>Heriot-Watt University, UK

PAGE 1216  
WedSe4.O4-7

**Control of Prosodic Focus in Corpus-Based Generation of Fundamental Frequency Based on the Generation Process Model**  
*Keiko Ochi, Keikichi Hirose, Nobuaki Minematsu, University of Tokyo, Japan*

---

---

## WedSe4.P1 : Automatic Speech Recognition: Adaptation I

Mezzanine Level Area A1, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Hervé Bourlard

---

PAGE 1217  
WedSe4.P1-1

### **Discriminative Training of Narrow Band - Wide Band Adapted Systems for Meeting Recognition**

*Martin Karafiát<sup>1</sup>, Lukáš Burget<sup>1</sup>, Thomas Hain<sup>2</sup>, Jan Černocký<sup>1</sup>*

<sup>1</sup>Brno University of Technology, Czech Republic; <sup>2</sup>University of Sheffield, UK

PAGE 1221  
WedSe4.P1-2

### **A Fast Speaker Adaptation Method Using Aspect Model**

*Seongjun Hahm<sup>1</sup>, Akinori Ito<sup>1</sup>, Shozo Makino<sup>1</sup>, Motoyuki Suzuki<sup>2</sup>*

<sup>1</sup>Tohoku University, Japan; <sup>2</sup>University of Tokushima, Japan

PAGE 1225  
WedSe4.P1-3

### **Probabilistic Latent Speaker Training for Large Vocabulary Speech Recognition**

*Dan Su, Xihong Wu, Huisheng Chi, Peking University, China*

PAGE 1229  
WedSe4.P1-4

### **Improvement of Eigenvoice-Based Speaker Adaptation by Parameter Space Clustering**

*Shutaro Tanji<sup>1</sup>, Koichi Shinoda<sup>1</sup>, Sadaoki Furui<sup>1</sup>, Antonio Ortega<sup>2</sup>*

<sup>1</sup>Tokyo Institute of Technology, Japan; <sup>2</sup>University of Southern California, USA

PAGE 1233  
WedSe4.P1-5

### **Study of Jacobian Compensation Using Linear Transformation of Conventional MFCC for VTLN**

*D.R. Sanand, S. Umesh, IIT Kanpur, India*

PAGE 1237  
WedSe4.P1-6

### **Adaptive HMM Topology for Speech Recognition**

*Chuan-Wei Ting, Kuo-Yuan Lee, Jen-Tzung Chien, National Cheng Kung University, Taiwan*

---

*WedSe4.P1 continued ...*

PAGE 1241  
WedSe4.P1-7

**Minimum Phone Error Discriminative Training for Mandarin Chinese Speaker Adaptation**

*Liang-Yu Chen*<sup>1</sup>, *Chun-Jen Lee*<sup>2</sup>, *Jyh-Shing Roger Jang*<sup>1</sup>

<sup>1</sup>*National Tsing Hua University, Taiwan*; <sup>2</sup>*Chunghwa Telecom Co. Ltd., Taiwan*

PAGE 1245  
WedSe4.P1-8

**Fast Speaker Adaptive Training for Speech Recognition**

*Daniel Povey, Hong-Kwang Jeff Kuo, Hagen Soltau, IBM T.J. Watson Research Center, USA*

---

---

## WedSe4.P2: Robust Automatic Speech Recognition II

Mezzanine Level Area A2, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Satoshi Nakamura

---

PAGE 1249  
WedSe4.P2-1

**Eigen-MLLR Environment/Speaker Compensation for Robust Speech Recognition**  
*Yuan-Fu Liao, Hung-Hsiang Fang, Chi-Hui Hsu, National Taipei University of Technology, Taiwan*

PAGE 1253  
WedSe4.P2-2

**Parameter Clustering and Sharing in Variable-Parameter HMMs for Noise Robust Speech Recognition**  
*Dong Yu<sup>1</sup>, Li Deng<sup>1</sup>, Yifan Gong<sup>2</sup>, Alex Acero<sup>1</sup>*  
<sup>1</sup>Microsoft Research, USA; <sup>2</sup>Microsoft Corporation, USA

PAGE 1257  
WedSe4.P2-3

**A Feature Compensation Approach Using High-Order Vector Taylor Series Approximation of an Explicit Distortion Model for Noisy Speech Recognition**  
*Jun Du, Qiang Huo, Microsoft Research Asia, China*

PAGE 1261  
WedSe4.P2-4

**N-Best Based Stochastic Mapping on Stereo HMM for Noise Robust Speech Recognition**  
*Xiaodong Cui<sup>1</sup>, Mohamed Afify<sup>2</sup>, Yuqing Gao<sup>1</sup>*  
<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>Orange Labs, Egypt

PAGE 1265  
WedSe4.P2-5

**Improving the Ensemble Speaker and Speaking Environment Modeling Approach by Enhancing the Precision of the Online Estimation Process**  
*Yu Tsao, Chin-Hui Lee, Georgia Institute of Technology, USA*

PAGE 1269  
WedSe4.P2-6

**Combining Noise Compensation and Missing-Feature Decoding for Large Vocabulary Speech Recognition in Noise**  
*Jianhua Lu, Ji Ming, Roger Woods, Queen's University Belfast, UK*



---

*WedSe4.P2 continued ...*

PAGE 1273  
WedSe4.P2-7

**Joint Bayesian Predictive Classification and Parallel Model Combination with Prior Scaling for Robust ASR**

*Svein Gunnar Pettersen, NTNU, Norway*

PAGE 1277  
WedSe4.P2-8

**Environment Mismatch Compensation Using Average Eigenspace for Speech Recognition**

*Abhishek Kumar, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1281  
WedSe4.P2-9

**Monte Carlo Model-Space Noise Adaptation for Speech Recognition**

*Daniel Povey, Brian Kingsbury, IBM T.J. Watson Research Center, USA*

PAGE 1285  
WedSe4.P2-10

**A 'Speechiness' Measure to Improve Speech Decoding in the Presence of Other Sound Sources**

*Ning Ma, Phil Green, University of Sheffield, UK*

PAGE 1289  
WedSe4.P2-11

**Feature Vector Normalization with Combined Standard and Throat Microphones for Robust ASR**

*Luis Buera, Antonio Miguel, Óscar Saz, Alfonso Ortega, Eduardo Lleida, Universidad de Zaragoza, Spain*

PAGE 1293  
WedSe4.P2-12

**Phone-Duration-Dependent Long-Term Dynamic Features for a Stochastic Model-Based Voice Activity Detection**

*Takashi Fukuda, Osamu Ichikawa, Masafumi Nishimura, IBM Japan Ltd., Japan*

PAGE 1297  
WedSe4.P2-13

**An On-Line Adaptation Technique for Emotional Speech Recognition Using Style Estimation with Multiple-Regression HMM**

*Yusuke Ijima, Makoto Tachibana, Takashi Nose, Takao Kobayashi, Tokyo Institute of Technology, Japan*

---

*WedSe4.P2 continued ...*

PAGE 1301  
WedSe4.P2-14

**HMM Adaptation Using Statistical Linear Approximation for Robust Automatic Speech Recognition**

*Michael Berkovitch, Ilan D. Shallom, Ben-Gurion University of the Negev, Israel*

PAGE 1305  
WedSe4.P2-15

**Beyond Linear Transforms: Efficient Non-Linear Dynamic Adaptation for Noise Robust Speech Recognition**

*Steven J. Rennie, Pierre L. Dognin, IBM T.J. Watson Research Center, USA*

PAGE 1309  
WedSe4.P2-16

**Rapid Unsupervised Speaker Adaptation Robust in Reverberant Environment Conditions**

*Randy Gomez<sup>1</sup>, Jani Even<sup>2</sup>, Kiyohiro Shikano<sup>2</sup>*

*<sup>1</sup>Kyoto University, Japan; <sup>2</sup>NAIST, Japan*

---

---

## WedSe4.P3: Features for Speech and Speaker Recognition

Mezzanine Level Area B3, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Richard M. Stern

---

PAGE 1313  
WedSe4.P3-1

### **Speaker Identification for Whispered Speech Based on Frequency Warping and Score Competition**

*Xing Fan, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1317  
WedSe4.P3-2

### **Experimental Evaluation of Multi-Band Position-Pitch Estimation (M-PoPi) Algorithm for Multi-Speaker Localization**

*Tania Habib, Lukas Ottowitz, Marián Képesi, Graz University of Technology, Austria*

PAGE 1321  
WedSe4.P3-3

### **Features for Automatic Detection of Voice Bars in Continuous Speech**

*Dhananjaya N.<sup>1</sup>, Rajendran S.<sup>2</sup>, B. Yegnanarayana<sup>2</sup>*

*<sup>1</sup>IIT Madras, India; <sup>2</sup>IIT Hyderabad, India*

PAGE 1325  
WedSe4.P3-4

### **Speaker Orientation Estimation Based on Hybridation of GCC-PHAT and HLBR**

*Carlos Segura<sup>1</sup>, Alberto Abad<sup>2</sup>, Javier Hernando<sup>1</sup>, Climent Nadeu<sup>1</sup>*

*<sup>1</sup>Universitat Politècnica de Catalunya, Spain; <sup>2</sup>INESC-ID/IST, Portugal*

PAGE 1329  
WedSe4.P3-5

### **Parallel and Hierarchical Speech Feature Classification Using Frame and Segment-Based Methods**

*Jun Hou, Lawrence Rabiner, Sorin Dusan, Rutgers University, USA*

PAGE 1333  
WedSe4.P3-6

### **Automatically Learning Speaker-Independent Acoustic Subword Units**

*Balakrishnan Varadarajan, Sanjeev Khudanpur, Johns Hopkins University, USA*

PAGE 1337  
WedSe4.P3-7

### **Human-Like Ears versus Two-Microphone Array, Which Works Better for Speaker Identification?**

*Waleed H. Abdulla, Yushi Zhang, University of Auckland, New Zealand*

---

*WedSe4.P3 continued ...*

PAGE 1341  
WedSe4.P3-8

**Is a Speech Recognizer Useful for Characteristic Analysis of Classroom Lecture Speech?**

*Kenji Kobayashi, Mitsuhiro Somiya, Hiromitsu Nishizaki, Yoshihiro Sekiguchi, University of Yamanashi, Japan*

PAGE 1345  
WedSe4.P3-9

**An Intuitive Class Discriminability Measure for Feature Selection in a Speech Recognition System**

*Ladan Golipour, Douglas O'Shaughnessy, Université du Québec, Canada*

PAGE 1349  
WedSe4.P3-10

**$f$ -Divergence is a Generalized Invariant Measure Between Distributions**

*Yu Qiao, Nobuaki Minematsu, University of Tokyo, Japan*

PAGE 1353  
WedSe4.P3-11

**Sparse Linear Predictors for Speech Processing**

*Daniele Giacobello<sup>1</sup>, Mads Græsbøll Christensen<sup>1</sup>, Joachim Dahl<sup>1</sup>, Søren Holdt Jensen<sup>1</sup>, Marc Moonen<sup>2</sup>*

*<sup>1</sup>Aalborg University, Denmark; <sup>2</sup>Katholieke Universiteit Leuven, Belgium*

PAGE 1357  
WedSe4.P3-12

**Frequency-Domain Parameter Estimations for Binary Masked Signals**

*J.X. Zhang<sup>1</sup>, Mads Græsbøll Christensen<sup>1</sup>, Joachim Dahl<sup>1</sup>, Søren Holdt Jensen<sup>1</sup>, Marc Moonen<sup>2</sup>*

*<sup>1</sup>Aalborg University, Denmark; <sup>2</sup>Katholieke Universiteit Leuven, Belgium*

PAGE 1361  
WedSe4.P3-13

**Decomposition of Rotational Distortion Caused by VTL Difference Using Eigenvalues of Its Transformation Matrix**

*Daisuke Saito, Nobuaki Minematsu, Keikichi Hirose, University of Tokyo, Japan*

PAGE 1365  
WedSe4.P3-14

**Region-Based Vocal Tract Length Normalization for ASR**

*Michail G. Maragakis, Alexandros Potamianos, Technical University of Crete, Greece*

---

---

## WedSe4.P4: Speaker Recognition: Kernel-Based and Session Mismatch

Mezzanine Level Area B4, Time 16:00 - 18:00, Wednesday 24th September 2008

Chair: Eliathamby Ambikairajah

---

PAGE 1369  
WedSe4.P4-1

### **Speaker Verification with Non-Audible Murmur Segments by Combining Global Alignment Kernel and Penalized Logistic Regression Machine**

*Hideki Okamoto*<sup>1</sup>, *Tomoko Matsui*<sup>2</sup>, *Hiromichi Kawanami*<sup>1</sup>, *Hiroshi Saruwatari*<sup>1</sup>, *Kiyohiro Shikano*<sup>1</sup>

<sup>1</sup>NAIST, Japan; <sup>2</sup>ISM, Japan

PAGE 1373  
WedSe4.P4-2

### **Analysis of Subspace Within-Class Covariance Normalization for SVM-Based Speaker Verification**

*Liang Lu*<sup>1</sup>, *Yuan Dong*<sup>1</sup>, *Xianyu Zhao*<sup>2</sup>, *Jian Zhao*<sup>1</sup>, *Chengyu Dong*<sup>2</sup>, *Haila Wang*<sup>2</sup>

<sup>1</sup>Beijing University of Posts & Telecommunications, China; <sup>2</sup>Orange Labs, China

PAGE 1377  
WedSe4.P4-3

### **Comparison of Input and Feature Space Nonlinear Kernel Nuisance Attribute Projections for Speaker Verification**

*Xianyu Zhao*<sup>1</sup>, *Yuan Dong*<sup>1</sup>, *Jian Zhao*<sup>2</sup>, *Liang Lu*<sup>2</sup>, *Jiqing Liu*<sup>2</sup>, *Haila Wang*<sup>1</sup>

<sup>1</sup>Orange Labs, China; <sup>2</sup>Beijing University of Posts & Telecommunications, China

PAGE 1381  
WedSe4.P4-4

### **A Generalised Derivative Kernel for Speaker Verification**

*C. Longworth*, *M.J.F. Gales*, University of Cambridge, UK

PAGE 1385  
WedSe4.P4-5

### **Modeling Prior Belief for Speaker Verification SVM Systems**

*Luciana Ferrer*, Stanford University, USA

PAGE 1389  
WedSe4.P4-6

### **Convergence Between SVM-Based and Distance-Based Paradigms for Speaker Recognition**

*Delphine Charlet*<sup>1</sup>, *Xianyu Zhao*<sup>2</sup>, *Yuan Dong*<sup>2</sup>

<sup>1</sup>Orange Labs, France; <sup>2</sup>Orange Labs, China

---

*WedSe4.P4 continued ...*

PAGE 1393  
WedSe4.P4-7

**High-Level Speaker Verification via Articulatory-Feature Based Sequence Kernels and SVM**

*Shi-Xiong Zhang, Man-Wai Mak, Hong Kong Polytechnic University, China*

PAGE 1397  
WedSe4.P4-8

**Characterizing Speech Utterances for Speaker Verification with Sequence Kernel SVM**

*Kong-Aik Lee<sup>1</sup>, Changhuai You<sup>1</sup>, Haizhou Li<sup>1</sup>, Tomi Kinnunen<sup>2</sup>, Donglai Zhu<sup>1</sup>*  
*<sup>1</sup>Institute for Infocomm Research, Singapore; <sup>2</sup>University of Joensuu, Finland*

PAGE 1401  
WedSe4.P4-9

**Development of the Primary CRIM System for the NIST 2008 Speaker Recognition Evaluation**

*Patrick Kenny, Najim Dehak, Pierre Ouellet, Vishwa Gupta, Pierre Dumouchel, CRIM, Canada*

PAGE 1405  
WedSe4.P4-10

**Making Confident Speaker Verification Decisions with Minimal Speech**

*Robbie Vogt, Sridha Sridharan, Michael Mason, Queensland University of Technology, Australia*

PAGE 1409  
WedSe4.P4-11

**Parallelized Factor Analysis and Feature Normalization for Automatic Speaker Verification**

*Jun Luo, Cheung-Chi Leung, Marc Ferràs, Claude Barras, LIMSI, France*

PAGE 1413  
WedSe4.P4-12

**Intersession Variability in Speaker Recognition: A Behind the Scene Analysis**

*Daniel Garcia-Romero, Carol Y. Espy-Wilson, University of Maryland, USA*

PAGE 1417  
WedSe4.P4-13

**Speaker Recognition Based on Variational Bayesian Method**

*Tatsuya Ito, Kei Hashimoto, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, Nagoya Institute of Technology, Japan*

---

*WedSe4.P4 continued ...*

PAGE 1421  
WedSe4.P4-14

**Factor Analysis Multi-Session Training Constraint in Session Compensation for Speaker Verification**

*Driss Matrouf, Jean-François Bonastre, Salah Eddine Mezaache, LIA, France*

PAGE 1425  
WedSe4.P4-15

**The Role of 'Delta' Features in Speaker Verification**

*Ying Liu, Martin J. Russell, Michael J. Carey, University of Birmingham, UK*

---

---

## ThuSe2.01: Broadcast Transcription Systems

Great Hall, Time 10:00 - 12:00, Thursday 25th September 2008

Chair: Tanja Schultz

---

PAGE 1429  
ThuSe2.01-1  
10:00 - 10:20

### **Investigating Morphological Decomposition for Transcription of Arabic Broadcast News and Broadcast Conversation Data**

*Lori Lamel, Abdel. Messaoudi, Jean-Luc Gauvain, LIMSI, France*

PAGE 1433  
ThuSe2.01-2  
10:20 - 10:40

### **Transcribing Broadcast Data Using MLP Features**

*Petr Fousek, Lori Lamel, Jean-Luc Gauvain, LIMSI, France*

PAGE 1437  
ThuSe2.01-3  
10:40 - 11:00

### **Development of the SRI/Nightingale Arabic ASR System**

*D. Vergyri<sup>1</sup>, A. Mandal<sup>1</sup>, Wen Wang<sup>1</sup>, Andreas Stolcke<sup>1</sup>, Jing Zheng<sup>1</sup>, Martin Graciarena<sup>1</sup>, D. Rybach<sup>2</sup>, Christian Gollan<sup>2</sup>, Ralf Schlüter<sup>2</sup>, Katrin Kirchhoff<sup>3</sup>, A. Faria<sup>4</sup>, Nelson Morgan<sup>4</sup>*

*<sup>1</sup>SRI International, USA; <sup>2</sup>RWTH Aachen University, Germany; <sup>3</sup>University of Washington, USA; <sup>4</sup>ICSI, USA*

PAGE 1441  
ThuSe2.01-4  
11:00 - 11:20

### **Towards Automatic Learning in LVCSR: Rapid Development of a Persian Broadcast Transcription System**

*Christian Gollan, Hermann Ney, RWTH Aachen University, Germany*

PAGE 1445  
ThuSe2.01-5  
11:20 - 11:40

### **The CMU-InterACT 2008 Mandarin Transcription System**

*Roger Hsiao, Mark Fuhs, Yik-Cheung Tam, Qin Jin, Tanja Schultz, Carnegie Mellon University, USA*

PAGE 1449  
ThuSe2.01-6  
11:40 - 12:00

### **Decoding-Time Prediction of Non-Verbalized Punctuation**

*Anoop Deoras<sup>1</sup>, Jürgen Fritsch<sup>2</sup>*

*<sup>1</sup>Johns Hopkins University, USA; <sup>2</sup>M\*Modal, USA*



---

---

## ThuSe2.O2: Voice Conversion and Modification

Plaza 1, Time 10:00 - 11:40, Thursday 25th September 2008

Chair: Chris Davis

---

PAGE 1453  
ThuSe2.O2-1  
10:00 - 10:20

### **On the Impact of Alignment on Voice Conversion Performance**

*Elina Helander*<sup>1</sup>, *Jan Schwarz*<sup>2</sup>, *Jani Nurminen*<sup>3</sup>, *Hanna Silen*<sup>1</sup>, *Moncef Gabbouj*<sup>1</sup>

<sup>1</sup>Tampere University of Technology, Finland; <sup>2</sup>Christian-Albrechts-Universität zu Kiel, Germany; <sup>3</sup>Nokia Devices R&D, Finland

PAGE 1457  
ThuSe2.O2-2  
10:20 - 10:40

### **The Linear Transformation of LF Glottal Waveforms for Voice Conversion**

*Arantza del Pozo*, *Steve Young*, University of Cambridge, UK

PAGE 1461  
ThuSe2.O2-3  
10:40 - 11:00

### **Maximum *a posteriori* Adaptation for Many-to-One Eigenvoice Conversion**

*Daisuke Tani*, *Tomoki Toda*, *Yamato Ohtani*, *Hiroshi Saruwatari*, *Kiyohiro Shikano*, NAIST, Japan

PAGE 1465  
ThuSe2.O2-4  
11:00 - 11:20

### **Improvement to a NAM Captured Whisper-to-Speech System**

*Viet-Anh Tran*, *Gérard Bailly*, *Hélène Lævenbruck*, *Christian Jutten*, GIPSA, France

PAGE 1469  
ThuSe2.O2-5  
11:20 - 11:40

### **Speaker Identification in Noise Mismatch Conditions Based on Jump Function Kolmogorov Analysis in Wavelet Domain**

*Huy Dat Tran*, *Haizhou Li*, Institute for Infocomm Research, Singapore

---

---

## ThuSe2.O3: Phonetics: General

Plaza 2, Time 10:00 - 11:40, Thursday 25th September 2008

Chair: Andrew Butcher

---

PAGE 1473  
ThuSe2.O3-1  
10:00 - 10:20

**Modelling Fine-Phonetic Detail in a Computational Model of Word Recognition**  
*Odette Scharenborg, Radboud Universiteit Nijmegen, The Netherlands*

PAGE 1477  
ThuSe2.O3-2  
10:20 - 10:40

**Pronunciation Reduction: How It Relates to Speech Style, Gender, and Age**  
*Helmer Strik, Joost van Doremalen, Catia Cucchiarini, Radboud Universiteit Nijmegen, The Netherlands*

PAGE 1481  
ThuSe2.O3-3  
10:40 - 11:00

**Analysis of Glottal Stops in Speech Signals**  
*B. Yegnanarayana<sup>1</sup>, Rajendran S.<sup>1</sup>, Hussien Seid Worku<sup>1</sup>, Dhananjaya N.<sup>2</sup>*  
<sup>1</sup>IIT Hyderabad, India; <sup>2</sup>IIT Madras, India

PAGE 1485  
ThuSe2.O3-4  
11:00 - 11:20

**The Acoustic to Articulation Mapping: Non-Linear or Non-Unique?**  
*Daniel Neiberg, G. Ananthkrishnan, Olov Engwall, KTH, Sweden*

PAGE 1489  
ThuSe2.O3-5  
11:20 - 11:40

**The Entropy of the Articulatory Phonological Code: Recognizing Gestures from Tract Variables**  
*Xiaodan Zhuang<sup>1</sup>, Hosung Nam<sup>2</sup>, Mark Hasegawa-Johnson<sup>1</sup>, Louis M. Goldstein<sup>2</sup>, Elliot Saltzman<sup>2</sup>*  
<sup>1</sup>University of Illinois at Urbana-Champaign, USA; <sup>2</sup>Haskins Laboratories, USA

---

---

## ThuSe2.O4: Special Session: Forensic Speaker Recognition — Traditional and Automatic Approaches

Plaza 3&4, Time 10:00 – 12:05, Thursday 25th September 2008

Chair: Geoffrey Stewart Morrison

---

PAGE 1493  
ThuSe2.O4-1  
10:00 – 10:25

### **Addressing Database Mismatch in Forensic Speaker Recognition with Ahumada III: A Public Real-Casework Database in Spanish**

*Daniel Ramos<sup>1</sup>, Joaquin Gonzalez-Rodriguez<sup>1</sup>, Javier Gonzalez-Dominguez<sup>1</sup>, Jose Juan Lucena-Molina<sup>2</sup>*

<sup>1</sup>Universidad Autónoma de Madrid, Spain; <sup>2</sup>Ministerio del Interior, Spain

PAGE 1497  
ThuSe2.O4-2  
10:25 – 10:50

### **FM Features for Automatic Forensic Speaker Recognition**

*Tharmarajah Thiruvaran, Eliathamby Ambikairajah, Julien Epps, University of New South Wales, Australia*

PAGE 1501  
ThuSe2.O4-3  
10:50 – 11:15

### **Automatic-Type Calibration of Traditionally Derived Likelihood Ratios: Forensic Analysis of Australian English /o/ Formant Trajectories**

*Geoffrey Stewart Morrison<sup>1</sup>, Yuko Kinoshita<sup>2</sup>*

<sup>1</sup>Australian National University, Australia; <sup>2</sup>University of Canberra, Australia

PAGE 1505  
ThuSe2.O4-4  
11:15 – 11:40

### **Forensic Speaker Verification Using Formant Features and Gaussian Mixture Models**

*Timo Becker<sup>1</sup>, Michael Jessen<sup>2</sup>, Catalin Grigoras<sup>3</sup>*

<sup>1</sup>Austrian Academy of Sciences, Austria; <sup>2</sup>Bundeskriminalamt, Germany; <sup>3</sup>Ministry of Justice, Romania

PAGE 1509  
ThuSe2.O4-5  
11:40 – 12:05

### **The Case for Automatic Higher-Level Features in Forensic Speaker Recognition**

*Elizabeth Shriberg, Andreas Stolcke, SRI International, USA*

---

---

## ThuSe2.P1 : Automatic Speech Recognition: Features I

Mezzanine Level Area A1, Time 10:00 - 12:00, Thursday 25th September 2008

Chair: Lin-shan Lee

---

PAGE 1513  
ThuSe2.P1-1

### **Group Delay Function for Improved Gender Identification**

*Kye-Hwan Lee, Sang-Ick Kang, Ji-Hyun Song, Joon-Hyuk Chang, Inha University, Korea*

PAGE 1517  
ThuSe2.P1-2

### **Frame-Synchronous and Local Confidence Measures for On-the-Fly Automatic Speech Recognition**

*Joseph Razik, Odile Mella, Dominique Fohr, Jean-Paul Haton, LORIA, France*

PAGE 1521  
ThuSe2.P1-3

### **Hilbert Envelope Based Spectro-Temporal Features for Phoneme Recognition in Telephone Speech**

*Samuel Thomas, Sriram Ganapathy, Hynek Hermansky, IDIAP Research Institute, Switzerland*

PAGE 1525  
ThuSe2.P1-4

### **Evidence of Coarticulation in a Phonological Feature Detection System**

*Abhijeet Sangwan, Ayako Ikeno, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1529  
ThuSe2.P1-5

### **Phoneme Recognition Based on Hybrid Neural Networks with Inhibition/Enhancement of Distinctive Phonetic Feature (DPF) Trajectories**

*Mohammad Nurul Huda, Kouichi Katsurada, Tsuneo Nitta, Toyohashi University of Technology, Japan*

PAGE 1533  
ThuSe2.P1-6

### **A Neural Network Based Nonlinear Feature Transformation for Speech Recognition**

*Hongbing Hu, Stephen A. Zahorian, Binghamton University, USA*

---

*ThuSe2.P1 continued ...*

PAGE 1537  
ThuSe2.P1-7

**Significance of Group Delay Based Acoustic Features in the Linguistic Search Space for Robust Speech Recognition**

*Ramya R.<sup>1</sup>, Rajesh M. Hegde<sup>2</sup>, Hema A. Murthy<sup>1</sup>*

*<sup>1</sup>IIT Madras, India; <sup>2</sup>IIT Kanpur, India*

PAGE 1541  
ThuSe2.P1-8

**Genetic Programming Based Optimization of Class-Dependent PCA for Extracting Robust MFCC**

*Houman Abbasian<sup>1</sup>, Babak Nasersharif<sup>2</sup>, Ahmad Akbari<sup>1</sup>*

*<sup>1</sup>Iran University of Science & Technology, Iran; <sup>2</sup>University of Guilan, Iran*

PAGE 1545  
ThuSe2.P1-9

**Comparison of AM-FM Based Features for Robust Speech Recognition**

*K.V.S. Narayana, T.V. Sreenivas, Indian Institute of Science, India*

PAGE 1549  
ThuSe2.P1-10

**Growing Bottleneck Features for Tandem ASR**

*Joe Frankel, Dong Wang, Simon King, University of Edinburgh, UK*

PAGE 1550  
ThuSe2.P1-11

**Landmark Based Recognition of Stops: Acoustic Attributes versus Smoothed Spectra**

*Veena Karjigi, Preeti Rao, IIT Bombay, India*

PAGE 1554  
ThuSe2.P1-12

**Speech Recognition Performance of CJLC: Corpus of Japanese Lecture Contents**

*Satoru Kogure<sup>1</sup>, Hiromitsu Nishizaki<sup>2</sup>, Masatoshi Tsuchiya<sup>3</sup>, Kazumasa Yamamoto<sup>3</sup>, Shingo Togashi<sup>3</sup>, Seiichi Nakagawa<sup>3</sup>*

*<sup>1</sup>Shizuoka University, Japan; <sup>2</sup>University of Yamanashi, Japan; <sup>3</sup>Toyohashi University of Technology, Japan*

---

---

## ThuSe2.P2 : Automatic Speech Recognition: Language Models II

Mezzanine Level Area A2, Time 10:00 - 12:00, Thursday 25th September 2008

Chair: Günther Ruske

---

PAGE 1558  
ThuSe2.P2-1

### **Evaluating Spoken Language Model Based on Filler Prediction Model in Speech Recognition**

*Kengo Ohta, Masatoshi Tsuchiya, Seiichi Nakagawa, Toyohashi University of Technology, Japan*

PAGE 1562  
ThuSe2.P2-2

### **Transcription-Less Call Routing Using Unsupervised Language Model Adaptation**

*Nicolae Duta, Nuance Communications, USA*

PAGE 1566  
ThuSe2.P2-3

### **Large Margin Multinomial Mixture Model for Text Categorization**

*Zhen-Yu Pan, Hui Jiang, York University, Canada*

PAGE 1570  
ThuSe2.P2-4

### **Language Modeling for Speech Recognition of Spoken Cantonese**

*Yu Ting Yeung, Houwei Cao, N.H. Zheng, Tan Lee, P.C. Ching, Chinese University of Hong Kong, China*

PAGE 1574  
ThuSe2.P2-5

### **Discriminative Rescoring Based on Minimization of Word Errors for Transcribing Broadcast News**

*Akio Kobayashi, Takahiro Oku, Shinichi Homma, Shoei Sato, Toru Imai, Tohru Takagi, NHK, Japan*

PAGE 1578  
ThuSe2.P2-6

### **Search and Classification Based Language Model Adaptation**

*Qin Shi<sup>1</sup>, Stephen M. Chu<sup>2</sup>, Wen Liu<sup>1</sup>, Hong-Kwang Jeff Kuo<sup>2</sup>, Yi Liu<sup>1</sup>, Yong Qin<sup>1</sup>*  
<sup>1</sup>IBM China Research Lab, China; <sup>2</sup>IBM T.J. Watson Research Center, USA

PAGE 1582  
ThuSe2.P2-7

### **Fast N-Gram Language Model Look-Ahead for Decoders with Static Pronunciation Prefix Trees**

*Marijn Huijbregts, Roeland Ordelman, Franciska de Jong, University of Twente, The Netherlands*

---

*ThuSe2.P2 continued ...*

PAGE 1586  
ThuSe2.P2-8

**Thai Named-Entity Recognition Using Class-Based Language Modeling on Multiple-Sized Subword Units**

*Kwanchiva Saykhum<sup>1</sup>, Vataya Boonpiam<sup>1</sup>, Nattanun Thatphithakkul<sup>1</sup>, Chai Wutiwiwatchai<sup>1</sup>, Cholwich Natthee<sup>2</sup>*

<sup>1</sup>NECTEC, Thailand; <sup>2</sup>Thammasat University, Thailand

PAGE 1590  
ThuSe2.P2-9

**Combining Statistical and Syntactical Systems for Spoken Language Understanding with Graphical Models**

*S. Schwärzler, J. Geiger, J. Schenk, M. Al-Hames, B. Hörnler, G. Ruske, Gerhard Rigoll, Technische Universität München, Germany*

PAGE 1594  
ThuSe2.P2-10

**Bag-of-Word Normalized N-Gram Models**

*Abhinav Sethy, Bhuvana Ramabhadran, IBM T.J. Watson Research Center, USA*

PAGE 1598  
ThuSe2.P2-11

**A Study of Unsupervised Clustering Techniques for Language Modeling**

*Sangyun Hahn<sup>1</sup>, Abhinav Sethy<sup>2</sup>, Hong-Kwang Jeff Kuo<sup>2</sup>, Bhuvana Ramabhadran<sup>2</sup>*

<sup>1</sup>University of Washington, USA; <sup>2</sup>IBM T.J. Watson Research Center, USA

PAGE 1602  
ThuSe2.P2-12

**Automatic Estimation of Language Model Parameters for Unseen Words Using Morpho-Syntactic Contextual Information**

*Ciro Martins<sup>1</sup>, António Teixeira<sup>1</sup>, João Neto<sup>2</sup>*

<sup>1</sup>Universidade de Aveiro, Portugal; <sup>2</sup>INESC-ID/IST, Portugal

PAGE 1606  
ThuSe2.P2-13

**Modeling the Effects on Time-into-Utterance on Word Probabilities**

*Nigel G. Ward, Alejandro Vega, University of Texas at El Paso, USA*

PAGE 1610  
ThuSe2.P2-14

**Inductive and Example-Based Learning for Text Classification**

*Ye-Yi Wang, Xiao Li, Alex Acero, Microsoft Research, USA*

---

*ThuSe2.P2 continued ...*

PAGE 1614  
ThuSe2.P2-15

**Comparing Word, Character, and Phoneme  $n$ -Grams for Subjective Utterance Recognition**

*Theresa Wilson<sup>1</sup>, Stephan Raaijmakers<sup>2</sup>*

<sup>1</sup>*University of Edinburgh, UK; <sup>2</sup>TNO ICT, The Netherlands*

PAGE 1618  
ThuSe2.P2-16

**IRSTLM: An Open Source Toolkit for Handling Large Scale Language Models**

*Marcello Federico, Nicola Bertoldi, Mauro Cettolo, FBK-irst, Italy*



---

---

## ThuSe2.P3 : Speech Resources and Technology Evaluation

Mezzanine Level Area B3, Time 10:00 – 12:00, Thursday 25th September 2008

Chair: Christoph Draxler

---

PAGE 1622  
ThuSe2.P3-1

### **Multi-Modal Recording, Analysis and Indexing of Poster Sessions**

*Tatsuya Kawahara*<sup>1</sup>, *Hisao Setoguchi*<sup>1</sup>, *Katsuya Takanashi*<sup>1</sup>, *Kentaro Ishizuka*<sup>2</sup>,  
*Shoko Araki*<sup>2</sup>

<sup>1</sup>Kyoto University, Japan; <sup>2</sup>NTT Corporation, Japan

PAGE 1626  
ThuSe2.P3-2

### **Automatic Pitch-Synchronous Phonetic Segmentation**

*Jindřich Matoušek*, *Jan Romportl*, University of West Bohemia, Czech Republic

PAGE 1630  
ThuSe2.P3-3

### **Two Protocols Comparing Human and Machine Phonetic Recognition Performance in Conversational Speech**

*Wade Shen*<sup>1</sup>, *Joseph Olive*<sup>2</sup>, *Douglas Jones*<sup>1</sup>

<sup>1</sup>MIT, USA; <sup>2</sup>DARPA, USA

PAGE 1634  
ThuSe2.P3-4

### **Analysis of Drivers' Speech in a Car Environment**

*Tomoyuki Kato*, *Jun Okamoto*, *Makoto Shozakai*, Asahi Kasei Corporation, Japan

PAGE 1638  
ThuSe2.P3-5

### **Preparing a Corpus of Dutch Spontaneous Dialogues for Automatic Phonetic Analysis**

*Barbara Schuppler*, *Mirjam Ernestus*, *Odette Scharenborg*, *Lou Boves*, Radboud Universiteit Nijmegen, The Netherlands

PAGE 1642  
ThuSe2.P3-6

### **Evaluation of Voice Activity and Voicing Detection**

*Bojan Kotnik*<sup>1</sup>, *Pierre Sendorek*<sup>2</sup>, *Sergey Astrov*<sup>3</sup>, *Turgay Koc*<sup>4</sup>, *Tolga Ciloglu*<sup>4</sup>,  
*Laura Docío Fernández*<sup>5</sup>, *Eduardo Rodríguez Banga*<sup>5</sup>, *Harald Höge*<sup>3</sup>, *Zdravko Kačič*<sup>1</sup>

<sup>1</sup>University of Maribor, Slovenia; <sup>2</sup>Telecom Paristech, France; <sup>3</sup>Siemens AG, Germany;  
<sup>4</sup>Middle East Technical University, Turkey; <sup>5</sup>University of Vigo, Spain

---

*ThuSe2.P3 continued ...*

PAGE 1646  
ThuSe2.P3-7

**WikiSpeech — A Content Management System for Speech Databases**  
*Christoph Draxler, Klaus Jänsch, LMU München, Germany*

PAGE 1650  
ThuSe2.P3-8

**Development and Evaluation of Polish Speech Corpus for Unit Selection Speech Synthesis Systems**  
*G. Demenko<sup>1</sup>, J. Bachan<sup>1</sup>, Bernd Möbius<sup>2</sup>, K. Klessa<sup>1</sup>, M. Szymański<sup>3</sup>, S. Grocholewski<sup>3</sup>*  
*<sup>1</sup>Adam Mickiewicz University, Poland; <sup>2</sup>University of Stuttgart, Germany; <sup>3</sup>Poznań University of Technology, Poland*

PAGE 1654  
ThuSe2.P3-9

**A Data Format Enabling Interoperation of Speech Recognition, Translation and Information Extraction Engines: The GALE Type System**  
*John F. Pitrelli, Burn L. Lewis, Edward A. Epstein, Jerome L. Quinn, Ganesh Ramaswamy, IBM T.J. Watson Research Center, USA*

PAGE 1658  
ThuSe2.P3-10

**A Rank-Predicted Pseudo-Greedy Approach to Efficient Text Selection from Large-Scale Corpus for Maximum Coverage of Target Units**  
*Wei Li<sup>1</sup>, Qiang Huo<sup>2</sup>*  
*<sup>1</sup>University of Hong Kong, China; <sup>2</sup>Microsoft Research Asia, China*

PAGE 1662  
ThuSe2.P3-11

**MeMo Workbench for Semi-Automated Usability Testing**  
*Klaus-Peter Engelbrecht<sup>1</sup>, Michael Kruppa<sup>2</sup>, Sebastian Möller<sup>1</sup>, Michael Quade<sup>1</sup>*  
*<sup>1</sup>Technische Universität Berlin, Germany; <sup>2</sup>DFKI GmbH, Germany*

PAGE 1666  
ThuSe2.P3-12

**MDS-Based Visualization Method for Multiple Speech Corpora**  
*Kimiko Yamakawa<sup>1</sup>, Tomoko Matsui<sup>2</sup>, Shuichi Itahashi<sup>1</sup>*  
*<sup>1</sup>NII, Japan; <sup>2</sup>ISM, Japan*

PAGE 1670  
ThuSe2.P3-13

**Scripted Dialogs versus Improvisation: Lessons Learned About Emotional Elicitation Techniques from the IEMOCAP Database**  
*Carlos Busso, Shrikanth S. Narayanan, University of Southern California, USA*

---

---

## ThuSe2.P4: Special Session: Prosody of Spontaneous Speech II

Mezzanine Level Area B4, Time 10:00 – 12:00, Thursday 25th September 2008

Chair: Hansjörg Mixdorff

---

PAGE 1674  
ThuSe2.P4-1

### **Analysis and Perception of Speech Under Physical Task Stress**

*Keith W. Godin, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1678  
ThuSe2.P4-2

### **An Analysis of Multimodal Cues of Interruption in Dyadic Spoken Interactions**

*Chi-Chun Lee, Sungbok Lee, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 1682  
ThuSe2.P4-3

### **Paralinguistic Effects on Turn-Taking Behavior in Expressive Conversation**

*Hiroki Mori, Hideki Kasuya, Utsunomiya University, Japan*

PAGE 1683  
ThuSe2.P4-4

### **Study on “Ng, A” Type of Discourse Markers in Standard Chinese**

*Zhigang Yin, Aijun Li, Ziyu Xiong, Chinese Academy of Social Sciences, China*

PAGE 1687  
ThuSe2.P4-5

### **How Can You Use Disfluencies and Still Sound as a Good Speaker?**

*Helena Moniz<sup>1</sup>, Ana Isabel Mata<sup>2</sup>, Isabel Trancoso<sup>3</sup>, M. Céu Viana<sup>2</sup>*

*<sup>1</sup>INESC-ID/CLUL, Portugal; <sup>2</sup>CLUL/FLUL, Portugal; <sup>3</sup>INESC-ID/IST, Portugal*

PAGE 1688  
ThuSe2.P4-6

### **What Makes a Good Speaker? Subject Ratings, Acoustic Measurements and Perceptual Evaluations**

*Eva Strangert<sup>1</sup>, Joakim Gustafson<sup>2</sup>*

*<sup>1</sup>Umeå University, Sweden; <sup>2</sup>KTH, Sweden*

PAGE 1692  
ThuSe2.P4-7

### **Towards Measuring Continuous Acoustic Feature Convergence in Unconstrained Spoken Dialogues**

*Spyros Kousidis, David Dorrán, Yi Wang, Brian Vaughan, Charlie Cullen, Dermot Campbell, Ciaran McDonnell, Eugene Coyle, Dublin Institute of Technology, Ireland*

*ThuSe2.P4 continued ...*

**Detection of Feeling Through Back-Channels in Spoken Dialogue**

*Tatsuya Kawahara*<sup>1</sup>, *Masayoshi Toyokura*<sup>1</sup>, *Teruhisa Misu*<sup>2</sup>, *Chiori Hori*<sup>2</sup>

<sup>1</sup>*Kyoto University, Japan*; <sup>2</sup>*NICT, Japan*

---

---

## ThuSe3.O1 : Automatic Speech Recognition: Adaptation II

Great Hall, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: Sadaoki Furui

---

PAGE 1697  
ThuSe3.O1-1  
13:30 - 13:50

### **Adaptive Training Using Discriminative Mapping Transforms**

*C.K. Raut, K. Yu, M.J.F. Gales, University of Cambridge, UK*

PAGE 1701  
ThuSe3.O1-2  
13:50 - 14:10

### **Speaker Adaptive Training Using Shift-MLLR**

*Jonas Löff, Christian Gollan, Hermann Ney, RWTH Aachen University, Germany*

PAGE 1705  
ThuSe3.O1-3  
14:10 - 14:30

### **XMLLR for Improved Speaker Adaptation in Speech Recognition**

*Daniel Povey, Hong-Kwang Jeff Kuo, IBM T.J. Watson Research Center, USA*

PAGE 1709  
ThuSe3.O1-4  
14:30 - 14:50

### **Effective Acoustic Adaptation for a Distant-Talking Interactive TV System**

*Jing Huang<sup>1</sup>, Mark Epstein<sup>1</sup>, Marco Matassoni<sup>2</sup>*

<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>FBK-irst, Italy

PAGE 1713  
ThuSe3.O1-5  
14:50 - 15:10

### **A Computationally Efficient Approach to Warp Factor Estimation in VTLN Using EM Algorithm and Sufficient Statistics**

*P.T. Akhil, S.P. Rath, S. Umesh, D.R. Sanand, IIT Kanpur, India*

PAGE 1717  
ThuSe3.O1-6  
15:10 - 15:30

### **A Reliable Technique for Detecting the Second Subglottal Resonance and Its Use in Cross-Language Speaker Adaptation**

*Shizhen Wang<sup>1</sup>, Steven M. Lulich<sup>2</sup>, Abeer Alwan<sup>1</sup>*

<sup>1</sup>University of California at Los Angeles, USA; <sup>2</sup>MIT, USA

---

---

## ThuSe3.O2: Applications in Education and Learning I

Plaza 1, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: Helmer Strik

---

PAGE 1721  
ThuSe3.O2-1  
13:30 - 13:50

### **Automatic Pronunciation Evaluation and Classification**

*Om D. Deshmukh, Sachindra Joshi, Ashish Verma, IBM India Research Lab, India*

PAGE 1725  
ThuSe3.O2-2  
13:50 - 14:10

### **Pronunciation Error Detection Techniques for Children's Speech**

*Daniel Bolaños, Wayne Ward, Barbara Wise, Sarel van Vuuren, University of Colorado at Boulder, USA*

PAGE 1729  
ThuSe3.O2-3  
14:10 - 14:30

### **Automatic Generation and Pruning of Phonetic Mispronunciations to Support Computer-Aided Pronunciation Training**

*Lan Wang<sup>1</sup>, Xin Feng<sup>1</sup>, Helen M. Meng<sup>2</sup>*

*<sup>1</sup>Chinese Academy of Sciences, China; <sup>2</sup>Chinese University of Hong Kong, China*

PAGE 1733  
ThuSe3.O2-4  
14:30 - 14:50

### **Automatic Children's Reading Tutor on Hand-Held Devices**

*Xiaolong Li, Li Deng, Yun-Cheng Ju, Alex Acero, Microsoft Research, USA*

PAGE 1737  
ThuSe3.O2-5  
14:50 - 15:10

### **A Japanese CALL System Based on Dynamic Question Generation and Error Prediction for ASR**

*Hongcui Wang, Tatsuya Kawahara, Kyoto University, Japan*

---

---

## ThuSe3.O3: Speech Pathologies

Plaza 2, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: Phil Green

---

PAGE 1741  
ThuSe3.O3-1  
13:30 - 13:50

### **Dysarthric Speech Database for Universal Access Research**

*Heejin Kim, Mark Hasegawa-Johnson, Adrienne Perlman, Jon Gunderson, Thomas S. Huang, Kenneth Watkin, Simone Frame, University of Illinois at Urbana-Champaign, USA*

PAGE 1745  
ThuSe3.O3-2  
13:50 - 14:10

### **Objective Intelligibility Assessment of Pathological Speakers**

*Catherine Middag<sup>1</sup>, Gwen Van Nuffelen<sup>2</sup>, Jean-Pierre Martens<sup>1</sup>, Marc De Bodt<sup>2</sup>*  
*<sup>1</sup>Ghent University, Belgium; <sup>2</sup>Antwerp University Hospital, Belgium*

PAGE 1749  
ThuSe3.O3-3  
14:10 - 14:30

### **Quantitative Analysis of Intonation Patterns Produced by Cantonese Speakers with Parkinson's Disease: A Preliminary Study**

*Joan K.-Y. Ma, Tara L. Whitehill, University of Hong Kong, China*

PAGE 1753  
ThuSe3.O3-4  
14:30 - 14:50

### **Phonetic-Acoustic and Feature Analyses by a Neural Network to Assess Speech Quality in Patients Treated for Head and Neck Cancer**

*Marieke de Bruijn<sup>1</sup>, Irma Verdonck de Leeuw<sup>1</sup>, Louis ten Bosch<sup>2</sup>, Joop Kuik<sup>1</sup>, Hugo Quené<sup>3</sup>, Lou Boves<sup>2</sup>, Hans Langendijk<sup>4</sup>, René Leemans<sup>1</sup>*

*<sup>1</sup>VU University Medical Center, The Netherlands; <sup>2</sup>Radboud Universiteit Nijmegen, The Netherlands; <sup>3</sup>Utrecht University, The Netherlands; <sup>4</sup>University Medical Center Groningen, The Netherlands*

PAGE 1757  
ThuSe3.O3-5  
14:50 - 15:10

### **Automatic Evaluation of Characteristic Speech Disorders in Children with Cleft Lip and Palate**

*Andreas Maier<sup>1</sup>, Florian Hönig<sup>1</sup>, Christian Hacker<sup>2</sup>, Maria Schuster<sup>3</sup>, Elmar Nöth<sup>1</sup>*

*<sup>1</sup>Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; <sup>2</sup>Elektrobit Automotive GmbH, Germany; <sup>3</sup>Universitätsklinikum Erlangen, Germany*

---

PAGE 1761  
ThuSe3.03-6  
15:10 - 15:30

*ThuSe3.03 continued ...*

**Application of Weighted Finite-State Transducers to Improve Recognition Accuracy for Dysarthric Speech**

*Omar Caballero Morales, Stephen Cox, University of East Anglia, UK*



---

---

## ThuSe3.04: Special Session: Consonant Challenge — Human-Machine Comparisons of Consonant Recognition in Noise

Plaza 3&4, Time 13:30 – 15:30, Thursday 25th September 2008

Chair: Odette Scharenborg and Martin Cooke

---

PAGE 1765  
ThuSe3.04-1

### **The Interspeech 2008 Consonant Challenge**

*Martin Cooke*<sup>1</sup>, *Odette Scharenborg*<sup>2</sup>

<sup>1</sup>University of Sheffield, UK; <sup>2</sup>Radboud Universiteit Nijmegen, The Netherlands

PAGE 1769  
ThuSe3.04-2

### **HMM-Based Estimation of Unreliable Spectral Components for Noise Robust Speech Recognition**

*Bengt J. Borgström*, *Abeer Alwan*, University of California at Los Angeles, USA

PAGE 1773  
ThuSe3.04-3

### **Gammatone-Domain Model Combination for Consonant Recognition in Noisy Environments**

*Jae Sam Yoon*, *Ji Hun Park*, *Hong Kook Kim*, GIST, Korea

PAGE 1777  
ThuSe3.04-4

### **On the Mask Modeling and Feature Representation in the Missing-Feature ASR: Evaluation on the Consonant Challenge**

*Peter Jančovič*, *Münevver Köküer*, University of Birmingham, UK

PAGE 1781  
ThuSe3.04-5

### **The Non-Native Consonant Challenge for European Languages**

*M. Luisa García Lecumberrri*<sup>1</sup>, *Martin Cooke*<sup>2</sup>, *Francesco Cutugno*<sup>3</sup>, *Mircea Giurgiu*<sup>4</sup>, *Bernd T. Meyer*<sup>5</sup>, *Odette Scharenborg*<sup>6</sup>, *Wim van Dommelen*<sup>7</sup>, *Jan Volin*<sup>8</sup>

<sup>1</sup>University of the Basque Country, Spain; <sup>2</sup>University of Sheffield, UK; <sup>3</sup>University of Naples “Federico II”, Italy; <sup>4</sup>Technical University of Cluj-Napoca, Romania; <sup>5</sup>Carl von Ossietzky Universität Oldenburg, Germany; <sup>6</sup>Radboud Universiteit Nijmegen, The Netherlands; <sup>7</sup>NTNU, Norway; <sup>8</sup>Charles University in Prague, Czech Republic

PAGE 1785  
ThuSe3.04-6

### **Noise Reduction Through Compressed Sensing**

*J.F. Gemmeke*, *B. Cranen*, Radboud Universiteit Nijmegen, The Netherlands

---

*ThuSe3.04 continued ...*

PAGE 1789  
ThuSe3.04-7

**Speech Recognition in Noisy Environments Using a Switching Linear Dynamic Model for Feature Enhancement**

*Björn Schuller<sup>1</sup>, Martin Wöllmer<sup>1</sup>, Tobias Moosmayr<sup>2</sup>, Gerhard Rigoll<sup>1</sup>*

*<sup>1</sup>Technische Universität München, Germany; <sup>2</sup>BMW Group, Germany*

PAGE 1793  
ThuSe3.04-8

**Improving Consonant Identification in Noise and Reverberation by Steady-State Suppression as a Preprocessing Approach**

*Nao Hodoshima, Wataru Yoshida, Takayuki Arai, Sophia University, Japan*

PAGE 1797  
ThuSe3.04-9

**Human Speech Perception and Feature Extraction**

*Bryce E. Lobdell, Mark Hasegawa-Johnson, Jont B. Allen, University of Illinois at Urbana-Champaign, USA*

---

---

## ThuSe3.P1 : Automatic Speech Recognition: Lexical and Prosodic Models

Mezzanine Level Area A1, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: Anders Lindström

---

PAGE 1801  
ThuSe3.P1-1

### **Improving Pronunciation Modeling for Non-Native Speech Recognition**

*Tien-Ping Tan, Laurent Besacier, LIG, France*

PAGE 1805  
ThuSe3.P1-2

### **Online Vocabulary Adaptation Using Contextual Information and Information Retrieval**

*Hagai Aronowitz, IBM Haifa Research Lab, Israel*

PAGE 1809  
ThuSe3.P1-3

### **Lexicon Expansion Using Pronunciation Variations Extracted on the Basis of Speaker-Related Deviation in Recognition Error Statistics**

*Yoshifumi Onishi, NEC Corporation, Japan*

PAGE 1813  
ThuSe3.P1-4

### **Better Nonnative Intonation Scores Through Prosodic Theory**

*Joseph Tepperman, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 1817  
ThuSe3.P1-5

### **Silence Models in Weighted Finite-State Transducers**

*Philip N. Garner, IDIAP Research Institute, Switzerland*

PAGE 1821  
ThuSe3.P1-6

### **Extracting Word-Pronunciation Pairs from Comparable Set of Text and Speech**

*Tetsuro Sasada, Shinsuke Mori, Tatsuya Kawahara, Kyoto University, Japan*

---

---

## ThuSe3.P2 : Speech Synthesis Methods II

Mezzanine Level Area A2, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: B. Yegnanarayana

---

PAGE 1825  
ThuSe3.P2-1

### **Synthesis by Generation and Concatenation of Multiform Segments**

*Vincent Pollet, Andrew Breen, Nuance Communications, Belgium*

PAGE 1829  
ThuSe3.P2-2

### **Glottal Spectral Separation for Parametric Speech Synthesis**

*João P. Cabral, Steve Renals, Korin Richmond, Junichi Yamagishi, University of Edinburgh, UK*

PAGE 1833  
ThuSe3.P2-3

### **Improving Speech Systems Built from Very Little Data**

*John Kominek, Sameer Badaskar, Tanja Schultz, Alan W. Black, Carnegie Mellon University, USA*

PAGE 1837  
ThuSe3.P2-4

### **Structure to Speech Conversion — Speech Generation Based on Infant-Like Vocal Imitation**

*Daisuke Saito, Satoshi Asakawa, Nobuaki Minematsu, Keikichi Hirose, University of Tokyo, Japan*

PAGE 1841  
ThuSe3.P2-5

### **Statistical Text-to-Speech Synthesis with Improved Dynamics**

*Stas Tiomkin, David Malah, Technion IIT, Israel*

PAGE 1845  
ThuSe3.P2-6

### **An Evaluation of Non-Standard Features for Grapheme-to-Phoneme Conversion**

*Gabriel Webster, Norbert Braunschweiler, Toshiba Research Europe Ltd., UK*

PAGE 1849  
ThuSe3.P2-7

### **Towards Flexible Speech Coding for Speech Synthesis: An LF + Modulated Noise Vocoder**

*Yannis Agiomyriannakis, Olivier Rosec, Orange Labs, France*

---

*ThuSe3.P2 continued ...*

PAGE 1853  
ThuSe3.P2-8

**Evaluation of Finnish Unit Selection and HMM-Based Speech Synthesis**

*Hanna Silen*<sup>1</sup>, *Elina Helander*<sup>1</sup>, *Jani Nurminen*<sup>2</sup>, *Moncef Gabbouj*<sup>1</sup>

<sup>1</sup>*Tampere University of Technology, Finland;* <sup>2</sup>*Nokia Devices R&D, Finland*

PAGE 1857  
ThuSe3.P2-9

**A Probabilistic Trajectory Synthesis System for Synthesising Visual Speech**

*Barry-John Theobald*, *Nicholas Wilkinson*, *University of East Anglia, UK*

PAGE 1861  
ThuSe3.P2-10

**Paralinguistic Elements in Speech Synthesis**

*Didier Cadic*<sup>1</sup>, *Lionel Segalen*<sup>2</sup>

<sup>1</sup>*Orange Labs, France;* <sup>2</sup>*Télécom Bretagne, France*

PAGE 1865  
ThuSe3.P2-11

**Building Sleek Synthesizers for Multi-Lingual Screen Reader**

*Veera Raghavendra E.*<sup>1</sup>, *B. Yegnanarayana*<sup>1</sup>, *Alan W. Black*<sup>2</sup>, *Kishore Prahallad*<sup>2</sup>

<sup>1</sup>*IIT Hyderabad, India;* <sup>2</sup>*Carnegie Mellon University, USA*

PAGE 1869  
ThuSe3.P2-12

**Unsupervised Adaptation for HMM-Based Speech Synthesis**

*Simon King*<sup>1</sup>, *Keiichi Tokuda*<sup>2</sup>, *Heiga Zen*<sup>2</sup>, *Junichi Yamagishi*<sup>1</sup>

<sup>1</sup>*University of Edinburgh, UK;* <sup>2</sup>*Nagoya Institute of Technology, Japan*

PAGE 1873  
ThuSe3.P2-13

**Investigating Festival's Target Cost Function Using Perceptual Experiments**

*Volker Strom*, *Simon King*, *University of Edinburgh, UK*

PAGE 1877  
ThuSe3.P2-14

**Modeling Austrian Dialect Varieties for TTS**

*Friedrich Neubarth*<sup>1</sup>, *Michael Pucher*<sup>2</sup>, *Christian Kranzler*<sup>2</sup>

<sup>1</sup>*OFAI, Austria;* <sup>2</sup>*ftw., Austria*

PAGE 1881  
ThuSe3.P2-15

**HMM-Based Finnish Text-to-Speech System Utilizing Glottal Inverse Filtering**

*Tuomo Raitio*<sup>1</sup>, *Antti Suni*<sup>2</sup>, *Hannu Pulakka*<sup>1</sup>, *Martti Vainio*<sup>2</sup>, *Paavo Alku*<sup>1</sup>

<sup>1</sup>*Helsinki University of Technology, Finland;* <sup>2</sup>*University of Helsinki, Finland*

---

*ThuSe3.P2 continued ...*

PAGE 1885  
ThuSe3.P2-16

**LTS Using Decision Forest of Regression Trees and Neural Networks**

*Tanuja Sarkar<sup>1</sup>, Sachin Joshi<sup>1</sup>, Sathish Chandra Pammi<sup>1</sup>, Kishore Prahallad<sup>2</sup>*

*<sup>1</sup>IIT Hyderabad, India; <sup>2</sup>Carnegie Mellon University, USA*

PAGE 1889  
ThuSe3.P2-17

**Automatic Word Stress Marking and Syllabification for Catalan TTS**

*Silvia Rustullet, Daniela Braga, João Nogueira, Miguel Sales Dias, Microsoft Language Development Center, Portugal*

---

---

## ThuSe3.P3 : Speaker Recognition: Adverse Conditions and Forensics

Mezzanine Level Area B3, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: Kiyohiro Shikano

---

PAGE 1893  
ThuSe3.P3-1

### **Robust Far-Field Speaker Identification Under Mismatched Conditions**

*Qin Jin, Tanja Schultz, Carnegie Mellon University, USA*

PAGE 1897  
ThuSe3.P3-2

### **Robust Speaker Verification Using Short-Time Frequency with Long-Time Window and Fusion of Multi-Resolutions**

*Chien-Lin Huang<sup>1</sup>, Bin Ma<sup>2</sup>, Chung-Hsien Wu<sup>1</sup>, Brian Mak<sup>3</sup>, Haizhou Li<sup>2</sup>*

*<sup>1</sup>National Cheng Kung University, Taiwan; <sup>2</sup>Institute for Infocomm Research, Singapore; <sup>3</sup>Hong Kong University of Science & Technology, China*

PAGE 1901  
ThuSe3.P3-3

### **Performance Improvement of Text-Independent Speaker Verification Systems Based on Histogram Enhancement in Noisy Environments**

*C.H. Kwon<sup>1</sup>, J.K. Choi<sup>2</sup>, Eliathamby Ambikairajah<sup>3</sup>*

*<sup>1</sup>Daejeon University, Korea; <sup>2</sup>BNSWorks, Korea; <sup>3</sup>University of New South Wales, Australia*

PAGE 1905  
ThuSe3.P3-4

### **Filling Acoustic Holes Through Leveraged Uncorellated GMMs for In-Set/Out-of-Set Speaker Recognition**

*Jun-Won Suh, Pongtep Angkititrakul, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1909  
ThuSe3.P3-5

### **Missing-Feature Method for Speaker Recognition in Band-Restricted Conditions**

*Wooil Kim, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 1913  
ThuSe3.P3-6

### **Robust Speaker Identification Using Cross-Correlation GTF-ICA Feature**

*Yushi Zhang, Waleed H. Abdulla, University of Auckland, New Zealand*

---

*ThuSe3.P3 continued ...*

PAGE 1917  
ThuSe3.P3-7

**Perceptual Speaker Identification Using Monosyllabic Stimuli — Effects of the Nucleus Vowels and Speaker Characteristics Contained in Nasals**

*Kanae Amino, Takayuki Arai, Sophia University, Japan*

PAGE 1921  
ThuSe3.P3-8

**Text-Dependent Speaker Recognition by Efficient Capture of Speaker Dynamics in Compressed Time-Frequency Representations of Speech**

*Amitava Das, Gokul Chittaranjan, Microsoft Research India, India*

PAGE 1925  
ThuSe3.P3-9

**Usefulness of Text-Conditioning and a New Database for Text-Dependent Speaker Recognition Research**

*Amitava Das, Gokul Chittaranjan, Gopala K. Anumanchipalli, Microsoft Research India, India*

PAGE 1929  
ThuSe3.P3-10

**Combination Method of Bone-Conduction Speech and Air-Conduction Speech for Speaker Recognition**

*Satoru Tsuge<sup>1</sup>, Takashi Osanai<sup>2</sup>, Hisanori Makinae<sup>2</sup>, Toshiaki Kamada<sup>2</sup>, Minoru Fukumi<sup>1</sup>, Shingo Kuroiwa<sup>3</sup>*

*<sup>1</sup>University of Tokushima, Japan; <sup>2</sup>National Research Institute of Police Science, Japan; <sup>3</sup>Chiba University, Japan*

PAGE 1933  
ThuSe3.P3-11

**MAP and Sub-Word Level T-Norm for Text-Dependent Speaker Recognition**

*Doroteo T. Toledano<sup>1</sup>, Daniel Hernandez-Lopez<sup>1</sup>, Cristina Esteve-Elizalde<sup>1</sup>, Joaquin Gonzalez-Rodriguez<sup>1</sup>, Ruben Fernandez Pozo<sup>2</sup>, Luis Hernandez Gomez<sup>2</sup>*

*<sup>1</sup>Universidad Autónoma de Madrid, Spain; <sup>2</sup>Universidad Politécnica de Madrid, Spain*

PAGE 1937  
ThuSe3.P3-12

**Forensic Speaker Recognition in Chinese: A Multivariate Likelihood Ratio Discrimination on /i/ and /y/**

*Cuiling Zhang<sup>1</sup>, Geoffrey Stewart Morrison<sup>2</sup>, Philip Rose<sup>2</sup>*

*<sup>1</sup>China Criminal Police University, China; <sup>2</sup>Australian National University, Australia*



---

*ThuSe3.P3 continued ...*

PAGE 1941  
ThuSe3.P3-13

**How Many Do We Need? Exploration of the Population Size Effect on the Performance of Forensic Speaker Classification**

*Shunichi Ishihara*<sup>1</sup>, *Yuko Kinoshita*<sup>2</sup>

<sup>1</sup>*Australian National University, Australia;* <sup>2</sup>*University of Canberra, Australia*

PAGE 1945  
ThuSe3.P3-14

**Comparing Prosodic Models for Speaker Recognition**

*Cheung-Chi Leung, Marc Ferràs, Claude Barras, Jean-Luc Gauvain, LIMSI, France*

PAGE 1949  
ThuSe3.P3-15

**Combination of Clean and Contaminated GMM/SVM for Far-Field Text-Independent Speaker Verification**

*Christian Zieger, Maurizio Omologo, FBK-irst, Italy*

---

---

## ThuSe3.P4 : Phonetics: Development, Learning, Cross-Language and Language-Specific

Mezzanine Level Area B4, Time 13:30 - 15:30, Thursday 25th September 2008

Chair: Sumio Ohno

---

PAGE 1953  
ThuSe3.P4-1

### **Language Identification on Code-Switching Utterances Using Multiple Cues**

*Bettina Braun*<sup>1</sup>, *Kristin Lemhöfer*<sup>2</sup>, *Anne Cutler*<sup>1</sup>

<sup>1</sup>Max Planck Institute for Psycholinguistics, The Netherlands; <sup>2</sup>Radboud Universiteit Nijmegen, The Netherlands

PAGE 1954  
ThuSe3.P4-2

### **The Strength of Stress-Related Lexical Competition Depends on the Presence of First-Syllable Stress**

*Eva Reinisch*, *Alexandra Jesse*, *James M. McQueen*, Max Planck Institute for Psycholinguistics, The Netherlands

PAGE 1955  
ThuSe3.P4-3

### **Word Stress Placement by Native Speakers and Japanese Learners of English**

*Keiichi Ishikawa*<sup>1</sup>, *Jun Nomura*<sup>2</sup>

<sup>1</sup>Kyoto Women's University, Japan; <sup>2</sup>University of Hawai'i at Mānoa, USA

PAGE 1959  
ThuSe3.P4-4

### **Schwa Variants in American English**

*H. Timothy Bunnell*<sup>1</sup>, *Jason Lilley*<sup>2</sup>

<sup>1</sup>Alfred I. duPont Hospital for Children, USA; <sup>2</sup>University of Delaware, USA

PAGE 1963  
ThuSe3.P4-5

### **Covariations of English Segmental Durations Across Speakers**

*Jiahong Yuan*, University of Pennsylvania, USA

PAGE 1964  
ThuSe3.P4-6

### **The Intelligibility of the English Vowel /ʌ/ Produced by Native Speakers of Japanese and Its Relations to the Acoustic Characteristics**

*Akiyo Joto*, Prefectural University of Hiroshima, Japan

PAGE 1968  
ThuSe3.P4-7

### **Rate Dependent Spectral Reduction for Voiceless Fricatives**

*Benjamin Weiss*, Technische Universität Berlin, Germany

---

*ThuSe3.P4 continued ...*

PAGE 1969  
ThuSe3.P4-8

**Investigating Perception of Places of Articulation in Sign and Speech**

*Stina Ojala*<sup>1</sup>, *Olli Aaltonen*<sup>2</sup>, *Tapio Salakoski*<sup>1</sup>

<sup>1</sup>*University of Turku, Finland*; <sup>2</sup>*University of Helsinki, Finland*

PAGE 1970  
ThuSe3.P4-9

**Six- and Twelve-Month-Olds' Discrimination of Native versus Non-Native Between- and Within-Organ Fricative Place Contrasts**

*Michael D. Tyler*<sup>1</sup>, *Catherine T. Best*<sup>1</sup>, *Louis M. Goldstein*<sup>2</sup>, *Mark Antoniou*<sup>1</sup>,  
*Lidija Krebs-Lazendic*<sup>1</sup>

<sup>1</sup>*University of Western Sydney, Australia*; <sup>2</sup>*Haskins Laboratories, USA*

PAGE 1971  
ThuSe3.P4-10

**"Your Baby Can't Hear You": How Mothers Talk to Infants with Simulated Hearing Loss**

*Christa Lam*, *Christine Kitamura*, *University of Western Sydney, Australia*

PAGE 1972  
ThuSe3.P4-11

**Development of Communicative Skills in 8- to 16-Month-Old Children: A Longitudinal Study**

*Eeva Klintfors*, *Ulla Sundberg*, *Francisco Lacerda*, *Ellen Marklund*, *Lisa Gustavsson*,  
*Ulla Bjursäter*, *Iris-Corinna Schwarz*, *Göran Söderlund*, *Stockholm University, Sweden*

PAGE 1976  
ThuSe3.P4-12

**Vocal Imitation in Early Language Acquisition**

*Lisa Gustavsson*, *Francisco Lacerda*, *Stockholm University, Sweden*

PAGE 1980  
ThuSe3.P4-13

**Computational Language Acquisition by Statistical Bottom-Up Processing**

*Okko Räsänen*, *Unto K. Laine*, *Toomas Altsaar*, *Helsinki University of Technology, Finland*

PAGE 1984  
ThuSe3.P4-14

**Lexical Analyses of Native and Non-Native English Language Instructor Speech Based on a Six-Month Co-Taught Classroom Video Corpus**

*Noriaki Katagiri*, *Goh Kawai*, *Hokkaido University, Japan*

*ThuSe3.P4 continued ...*

**Perception and Production of Consonant Clusters in Japanese-English Bilingual  
and Japanese Monolingual Speakers**

*Hinako Masuda, Takayuki Arai, Sophia University, Japan*

---

---

## ThuSe4.O1 : Robust Automatic Speech Recognition III

Great Hall, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Roberto Togneri

---

PAGE 1992  
ThuSe4.O1-1  
16:00 - 16:20

### **On a Generalization of Margin-Based Discriminative Training to Robust Speech Recognition**

*Jinyu Li, Chin-Hui Lee, Georgia Institute of Technology, USA*

PAGE 1996  
ThuSe4.O1-2  
16:20 - 16:40

### **Discriminative Classifiers with Generative Kernels for Noise Robust ASR**

*M.J.F. Gales, C. Longworth, University of Cambridge, UK*

PAGE 2000  
ThuSe4.O1-3  
16:40 - 17:00

### **Covariance Modelling for Noise-Robust Speech Recognition**

*R.C. van Dalen, M.J.F. Gales, University of Cambridge, UK*

PAGE 2004  
ThuSe4.O1-4  
17:00 - 17:20

### **Exploiting Spatial-Temporal Feature Distribution Characteristics for Robust Speech Recognition**

*Wei-Hau Chen, Shih-Hsiang Lin, Berlin Chen, National Taiwan Normal University, Taiwan*

PAGE 2008  
ThuSe4.O1-5  
17:20 - 17:40

### **Study of Integration of Statistical Model-Based Voice Activity Detection and Noise Suppression**

*Masakiyo Fujimoto, Kentaro Ishizuka, Tomohiro Nakatani, NTT Corporation, Japan*

PAGE 2012  
ThuSe4.O1-6  
17:40 - 18:00

### **Neural Network Based Regression for Robust Overlapping Speech Recognition Using Microphone Arrays**

*Weifeng Li, John Dines, Mathew Magimai Doss, Hervé Bouchard, IDIAP Research Institute, Switzerland*

---

---

## ThuSe4.O2: Multimodal Signal Processing

Plaza 1, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Roland Göcke

---

PAGE 2016  
ThuSe4.O2-1  
16:00 - 16:20

### **Lip Synchronization: From Phone Lattice to PCA Eigen-Projections Using Neural Networks**

*Samer Al Moubayed*<sup>1</sup>, *Michael De Smet*<sup>2</sup>, *Hugo Van hamme*<sup>2</sup>

<sup>1</sup>*KTH, Sweden*; <sup>2</sup>*Katholieke Universiteit Leuven, Belgium*

PAGE 2020  
ThuSe4.O2-2  
16:20 - 16:40

### **Building and Combining Document and Music Spaces for Music Query-by-Webpage System**

*Ryoei Takahashi*, *Yasunori Ohishi*, *Norihide Kitaoka*, *Kazuya Takeda*, *Nagoya University, Japan*

PAGE 2024  
ThuSe4.O2-3  
16:40 - 17:00

### **Improving Searching Speed and Accuracy of Query by Humming System Based on Three Methods: Feature Fusion, Candidates Set Reduction and Multiple Similarity Measurement Rescoring**

*Lei Wang*, *Shen Huang*, *Sheng Hu*, *Jiaen Liang*, *Bo Xu*, *Chinese Academy of Sciences, China*

PAGE 2028  
ThuSe4.O2-4  
17:00 - 17:20

### **Towards a Segmental Vocoder Driven by Ultrasound and Optical Images of the Tongue and Lips**

*Thomas Hueber*<sup>1</sup>, *Gérard Chollet*<sup>2</sup>, *Bruce Denby*<sup>3</sup>, *Gérard Dreyfus*<sup>1</sup>, *Maureen Stone*<sup>4</sup>

<sup>1</sup>*LE-ESPCI, France*; <sup>2</sup>*LTCl, France*; <sup>3</sup>*Université Pierre et Marie Curie, France*;

<sup>4</sup>*University of Maryland, USA*

PAGE 2032  
ThuSe4.O2-5  
17:20 - 17:40

### **Phone Recognition from Ultrasound and Optical Video Sequences for a Silent Speech Interface**

*Thomas Hueber*<sup>1</sup>, *Gérard Chollet*<sup>2</sup>, *Bruce Denby*<sup>3</sup>, *Gérard Dreyfus*<sup>1</sup>, *Maureen Stone*<sup>4</sup>

<sup>1</sup>*LE-ESPCI, France*; <sup>2</sup>*LTCl, France*; <sup>3</sup>*Université Pierre et Marie Curie, France*;

<sup>4</sup>*University of Maryland, USA*

---

PAGE 2036  
ThuSe4.O2-6  
17:40 - 18:00

*ThuSe4.O2 continued ...*

**Feature Space Transforms for Czech Sign-Language Recognition**

*Jan Trmal, Marek Hruží, Jan Zelinka, Pavel Campr, Luděk Müller, University of West  
Bohemia, Czech Republic*

---

---

## ThuSe4.O3: Speech Perception

Plaza 2, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Jeusun Kim

---

PAGE 2040  
ThuSe4.O3-1  
16:00 - 16:20

### **Masked Speech Priming: No priming in Dense Neighbourhoods**

*Chris Davis, Jeusun Kim, Angelo Barbaro, University of Western Sydney, Australia*

PAGE 2044  
ThuSe4.O3-2  
16:20 - 16:40

### **Integration of Audiovisual Speech and Priming Effects**

*Azra N. Ali, University of Huddersfield, UK*

PAGE 2048  
ThuSe4.O3-3  
16:40 - 17:00

### **Similarity Between Vowels Influences Response Execution in Word Identification**

*Jason D. Zevin, Thomas A. Farmer, Weill-Cornell Medical College, USA*

PAGE 2052  
ThuSe4.O3-4  
17:00 - 17:20

### **Phonotactically Well-Formed Onset Clusters as Processing Units in Word Recognition**

*Tom Lentz, Utrecht University, The Netherlands*

PAGE 2056  
ThuSe4.O3-5  
17:20 - 17:40

### **Prelexically-Driven Perceptual Retuning of Phoneme Boundaries**

*Anne Cutler<sup>1</sup>, James M. McQueen<sup>1</sup>, Sally Butterfield<sup>2</sup>, Dennis Norris<sup>2</sup>*

*<sup>1</sup>Max Planck Institute for Psycholinguistics, The Netherlands; <sup>2</sup>MRC Cognition & Brain Sciences Unit, UK*

PAGE 2057  
ThuSe4.O3-6  
17:40 - 18:00

### **Visual Speech Modifies the Phoneme Restoration Effect**

*Erin Cvejic, Jeusun Kim, Chris Davis, University of Western Sydney, Australia*



---

---

## ThuSe4.04: Evaluation and Standardisation of Spoken-Language Technology

Plaza 3&4, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Bruce Millar

---

PAGE 2058  
ThuSe4.04-1  
16:00 - 16:20

### **An Objective Singing Evaluation Approach by Relating Acoustic Measurements to Perceptual Ratings**

*Chuan Cao, Ming Li, Jian Liu, Yonghong Yan, Chinese Academy of Sciences, China*

PAGE 2062  
ThuSe4.04-2  
16:20 - 16:40

### **On the Perceived Quality of Noise Reduced Signals**

*Valérie Gautier-Turbin, Laetitia Gros, Orange Labs, France*

PAGE 2066  
ThuSe4.04-3  
16:40 - 17:00

### **A Methodology and Tool Suite for Evaluation of Accuracy of Interoperating Statistical Natural Language Processing Engines**

*Uma Murthy<sup>1</sup>, John F. Pitrelli<sup>2</sup>, Ganesh Ramaswamy<sup>2</sup>, Martin Franz<sup>2</sup>, Burn L. Lewis<sup>2</sup>*

*<sup>1</sup>Virginia Tech, USA; <sup>2</sup>IBM T.J. Watson Research Center, USA*

PAGE 2070  
ThuSe4.04-4  
17:00 - 17:20

### **An Empirical Analysis of Word Error Rate and Keyword Error Rate**

*Youngja Park<sup>1</sup>, Siddharth Patwardhan<sup>2</sup>, Karthik Visweswariah<sup>3</sup>, Stephen C. Gates<sup>1</sup>*

*<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>University of Utah, USA; <sup>3</sup>IBM India Research Lab, India*

PAGE 2074  
ThuSe4.04-5  
17:20 - 17:40

### **Measuring Speech Quality Impact on Tasks Performance**

*Virginie Durin, Laetitia Gros, Orange Labs, France*

PAGE 2078  
ThuSe4.04-6  
17:40 - 18:00

### **Voice Commands in Home Environment — A Consumer Survey**

*Hannu Soronen<sup>1</sup>, Markku Turunen<sup>2</sup>, Jaakko Hakulinen<sup>2</sup>*

*<sup>1</sup>Tampere University of Technology, Finland; <sup>2</sup>University of Tampere, Finland*

---

---

## ThuSe4.P1 : Automatic Speech Recognition: Search Methods

Mezzanine Level Area A1, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Hugo Van hamme

---

- PAGE 2082  
ThuSe4.P1-1      **Extended Partial Distance Elimination and Dynamic Gaussian Selection for Fast Likelihood Computation**  
*Ghazi Bouselmi, Jun Cai, LORIA, France*
- PAGE 2086  
ThuSe4.P1-2      **Improving the Multigram Algorithm by Using Lattices as Input**  
*Joris Driesen, Hugo Van hamme, Katholieke Universiteit Leuven, Belgium*
- PAGE 2090  
ThuSe4.P1-3      **Backward Viterbi Beam Search for Utilizing Dynamic Task Complexity Information**  
*Min Tang, Philippe Di Cristo, VoiceBox Technologies, USA*
- PAGE 2094  
ThuSe4.P1-4      **Fast Speech Decoding Through Phone Confusion Networks**  
*Nicola Bertoldi, Marcello Federico, Daniele Falavigna, Matteo Gerosa, FBK-irst, Italy*
- PAGE 2098  
ThuSe4.P1-5      **High-Performance Low-Latency Speech Recognition via Multi-Layered Feature Streaming and Fast Gaussian Computation**  
*Liang Gu, Jian Xue, Xiaodong Cui, Yuqing Gao, IBM T.J. Watson Research Center, USA*
- PAGE 2102  
ThuSe4.P1-6      **A Low-Power Hardware Search Architecture for Speech Recognition**  
*Patrick J. Bourke, Rob A. Rutenbar, Carnegie Mellon University, USA*
- PAGE 2106  
ThuSe4.P1-7      **Phonetic Query Expansion for Spoken Document Retrieval**  
*Jonathan Mamou<sup>1</sup>, Bhuvana Ramabhadran<sup>2</sup>*  
*<sup>1</sup>IBM Haifa Research Lab, Israel; <sup>2</sup>IBM T.J. Watson Research Center, USA*
- PAGE 2110  
ThuSe4.P1-8      **Implementation and Evaluation of Fast On-the-Fly WFST Composition Algorithms**  
*Tasuku Oonishi, Paul R. Dixon, Koji Iwano, Sadaoki Furui, Tokyo Institute of Technology, Japan*

---

---

## ThuSe4.P2 : Speech Synthesis: Prosody and Emotion I

Mezzanine Level Area A2, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Chiu-yu Tseng

---

PAGE 2114  
ThuSe4.P2-1

### **Analysis of Voice-Quality Features of Speech That Expresses “Anger”, “Joy”, and “Sadness” Uttered by Radio Actors and Actresses**

*Shoichi Takeda*<sup>1</sup>, *Yuuri Yasuda*<sup>2</sup>, *Risako Isobe*<sup>3</sup>, *Shogo Kiryu*<sup>3</sup>, *Makiko Tsuru*<sup>1</sup>

<sup>1</sup>Kinki University, Japan; <sup>2</sup>Osaka Gas Information System Research Institute Co. Ltd., Japan; <sup>3</sup>Musashi Institute of Technology, Japan

PAGE 2118  
ThuSe4.P2-2

### **Including Pitch Accent Optionality in Unit Selection Text-to-Speech Synthesis**

*Leonardo Badino*, *Robert A.J. Clark*, *Volker Strom*, University of Edinburgh, UK

PAGE 2122  
ThuSe4.P2-3

### **Emotion Conversion Using F0 Segment Selection**

*Zeynep Inanoglu*, *Steve Young*, University of Cambridge, UK

PAGE 2126  
ThuSe4.P2-4

### **Generating Natural F0 Trajectory with Additive Trees**

*Yao Qian*<sup>1</sup>, *Hui Liang*<sup>2</sup>, *Frank K. Soong*<sup>1</sup>

<sup>1</sup>Microsoft Research Asia, China; <sup>2</sup>Shanghai Jiao Tong University, China

PAGE 2130  
ThuSe4.P2-5

### **Generating Intonation from a Mixed CART-HMM Model for Speech Synthesis**

*Cédric Boidin*<sup>1</sup>, *Olivier Boeffard*<sup>2</sup>

<sup>1</sup>Orange Labs, France; <sup>2</sup>IRISA, France

PAGE 2134  
ThuSe4.P2-6

### **Intonation Modeling of Mandarin Chinese Using a Superpositional Approach**

*Pablo Daniel Agüero*<sup>1</sup>, *Antonio Bonafonte*<sup>2</sup>, *Lu Yu*<sup>2</sup>, *Juan Carlos Tulli*<sup>1</sup>

<sup>1</sup>University of Mar del Plata, Argentina; <sup>2</sup>Universitat Politècnica de Catalunya, Spain

PAGE 2138  
ThuSe4.P2-7

### **Two-Stage Prosody Prediction for Emotional Text-to-Speech Synthesis**

*Hao Tang*, *Xi Zhou*, *Matthias Odisio*, *Mark Hasegawa-Johnson*, *Thomas S. Huang*, University of Illinois at Urbana-Champaign, USA

*ThuSe4.P2 continued ...*

**Prosody Boundary Detection Through Context-Dependent Position Models**

*Yue-Ning Hu*<sup>1</sup>, *Min Chu*<sup>2</sup>, *Chao Huang*<sup>2</sup>, *Yan-Ning Zhang*<sup>1</sup>

<sup>1</sup>*Northwestern Polytechnical University, China;* <sup>2</sup>*Microsoft Research Asia, China*

---

---

## ThuSe4.P3 : Language Information Retrieval Systems

Mezzanine Level Area B3, Time 16:00 - 18:00, Thursday 25th September 2008

Chair: Piero Cosi

---

- PAGE 2146  
ThuSe4.P3-1      **Addressing the Out-of-Vocabulary Problem for Large-Scale Chinese Spoken Term Detection**  
*Sha Meng<sup>1</sup>, Jian Shao<sup>2</sup>, Roger Peng Yu<sup>2</sup>, Jia Liu<sup>1</sup>, Frank Seide<sup>2</sup>*  
<sup>1</sup>Tsinghua University, China; <sup>2</sup>Microsoft Research Asia, China
- PAGE 2150  
ThuSe4.P3-2      **Towards Vocabulary-Independent Speech Indexing for Large-Scale Repositories**  
*Jian Shao<sup>1</sup>, Roger Peng Yu<sup>1</sup>, Qingwei Zhao<sup>2</sup>, Yonghong Yan<sup>2</sup>, Frank Seide<sup>1</sup>*  
<sup>1</sup>Microsoft Research Asia, China; <sup>2</sup>Chinese Academy of Sciences, China
- PAGE 2154  
ThuSe4.P3-3      **Towards the Integration of Automatic Speech Recognition and Information Retrieval for Spoken Query Processing**  
*A. Moreno-Daniel<sup>1</sup>, J. Wilpon<sup>2</sup>, B.-H. Juang<sup>1</sup>, S. Parthasarathy<sup>2</sup>*  
<sup>1</sup>Georgia Institute of Technology, USA; <sup>2</sup>AT&T Labs Research, USA
- PAGE 2158  
ThuSe4.P3-4      **Reducing the Effect of OOV Query Words by Using Morph-Based Spoken Document Retrieval**  
*Ville T. Turunen, Helsinki University of Technology, Finland*
- PAGE 2162  
ThuSe4.P3-5      **Bayesian Latent Topic Clustering Model**  
*Meng-Sung Wu, Jen-Tzung Chien, National Cheng Kung University, Taiwan*
- PAGE 2166  
ThuSe4.P3-6      **Spoken Document Retrieval by Translating Recognition Candidates into Correct Transcriptions**  
*Tomoyosi Akiba, Yusuke Yokota, Toyohashi University of Technology, Japan*
- PAGE 2170  
ThuSe4.P3-7      **Audio Indexing for an Interactive Italian Literature Management System**  
*Carlo Drioli<sup>1</sup>, Piero Cosi<sup>2</sup>*  
<sup>1</sup>Università di Verona, Italy; <sup>2</sup>CNR-ISTC, Italy

---

*ThuSe4.P3 continued ...*

PAGE 2171  
ThuSe4.P3-8

**Open-Vocabulary Spoken-Document Retrieval Based on Query Expansion Using Related Web Documents**

*Makoto Terao, Takafumi Koshinaka, Shinichi Ando, Ryosuke Isotani, Akitoshi Okumura, NEC Corporation, Japan*

PAGE 2175  
ThuSe4.P3-9

**Discriminative Graph Training for Ultra-Fast Low-Footprint Speech Indexing**

*Upendra Chaudhari, Hong-Kwang Jeff Kuo, Brian Kingsbury, IBM T.J. Watson Research Center, USA*

PAGE 2179  
ThuSe4.P3-10

**A Language-Modeling Approach to Inverse Text Normalization and Data Cleanup for Multimodal Voice Search Applications**

*Yun-Cheng Ju<sup>1</sup>, Julian Odell<sup>2</sup>*

*<sup>1</sup>Microsoft Research, USA; <sup>2</sup>Microsoft Corporation, USA*

PAGE 2183  
ThuSe4.P3-11

**Topic Segmentation and Indexation in a Media Watch System**

*Rui Amaral, Isabel Trancoso, INESC-ID/IST, Portugal*

PAGE 2187  
ThuSe4.P3-12

**Vocabulary Independent Discriminative Term Frequency Estimation**

*J. Scott Olsson, University of Maryland, USA*

PAGE 2191  
ThuSe4.P3-13

**Spoken Keyword Spotting via Multi-Lattice Alignment**

*Hui Lin, Alex Stupakov, Jeff A. Bilmes, University of Washington, USA*

PAGE 2195  
ThuSe4.P3-14

**Robust Spoken Term Detection Using Combination of Phone-Based and Word-Based Recognition**

*Kenji Iwata, Koichi Shinoda, Sadaoki Furui, Tokyo Institute of Technology, Japan*

---

---

## ThuSe4.P4: Applications for the Aged and Handicapped

Mezzanine Level Area B4, Time 16:00 – 18:00, Thursday 25th September 2008

Chair: Stephen Cox

---

PAGE 2199  
ThuSe4.P4-1

### **Language Model Adaptation for a Speech to Sign Language Translation System Using Web Frequencies and a MAP Framework**

*Luis Fernando D'Haro<sup>1</sup>, Rubén San-Segundo<sup>1</sup>, Ricardo de Córdoba<sup>1</sup>, Jan Bungeroth<sup>2</sup>, Daniel Stein<sup>2</sup>, Hermann Ney<sup>2</sup>*

<sup>1</sup>Universidad Politécnica de Madrid, Spain; <sup>2</sup>RWTH Aachen University, Germany

PAGE 2203  
ThuSe4.P4-2

### **Hearing at Home — Communication Support in Home Environments for Hearing Impaired Persons**

*Jonas Beskow<sup>1</sup>, Björn Granström<sup>1</sup>, Peter Nordqvist<sup>1</sup>, Samer Al Moubayed<sup>1</sup>, Giampiero Salvi<sup>1</sup>, Tobias Herzke<sup>2</sup>, Arne Schulz<sup>3</sup>*

<sup>1</sup>KTH, Sweden; <sup>2</sup>HörTech, Germany; <sup>3</sup>OFFIS, Germany

PAGE 2207  
ThuSe4.P4-3

### **Traveling Wave Based Group Delays for Cochlear Implant Speech Processing**

*Daniel A. Taft, David B. Grayden, Anthony N. Burkitt, University of Melbourne, Australia*

PAGE 2208  
ThuSe4.P4-4

### **Multimodal Perception of Mandarin Tone for Cochlear Implant Users**

*Damien J. Smith, Denis Burnham, University of Western Sydney, Australia*

PAGE 2209  
ThuSe4.P4-5

### **Evaluation of Speaking-Aid System with Voice Conversion for Laryngectomees Toward Its Use in Practical Environments**

*Keigo Nakamura<sup>1</sup>, Tomoki Toda<sup>1</sup>, Yoshitaka Nakajima<sup>2</sup>, Hiroshi Saruwatari<sup>1</sup>, Kiyohiro Shikano<sup>1</sup>*

<sup>1</sup>NAIST, Japan; <sup>2</sup>Osaka University, Japan

---

*ThuSe4.P4 continued ...*

PAGE 2213  
ThuSe4.P4-6

**An Acoustic Typology of Apraxic Speech — Toward Reliable Diagnosis**

*Jacqueline McKechnie<sup>1</sup>, Kirrie J. Ballard<sup>1</sup>, Donald A. Robin<sup>2</sup>, Adam Jacks<sup>2</sup>, Sallyanne Palethorpe<sup>3</sup>, Kristin M. Rosen<sup>4</sup>*

<sup>1</sup>University of Sydney, Australia; <sup>2</sup>University of Texas Health Science Center at San Antonio, USA; <sup>3</sup>Macquarie University, Australia; <sup>4</sup>Independent Researcher, Australia

PAGE 2214  
ThuSe4.P4-7

**Dysphonic Voices and the 0–3000Hz Frequency Band**

*G. Pouchoulin<sup>1</sup>, C. Fredouille<sup>1</sup>, Jean-François Bonastre<sup>1</sup>, A. Ghio<sup>2</sup>, A. Giovanni<sup>2</sup>*

<sup>1</sup>LIA, France; <sup>2</sup>LPL, France

PAGE 2218  
ThuSe4.P4-8

**Verifying Pronunciation Accuracy from Speakers with Neuromuscular Disorders**

*Shou-Chun Yin<sup>1</sup>, Richard C. Rose<sup>1</sup>, Oscar Saz<sup>2</sup>, Eduardo Lleida<sup>2</sup>*

<sup>1</sup>McGill University, Canada; <sup>2</sup>Universidad de Zaragoza, Spain

PAGE 2222  
ThuSe4.P4-9

**Multi-Band and Multi-Cue Analyses of Disordered Connected Speech**

*A. Alpan<sup>1</sup>, Y. Maryn<sup>2</sup>, F. Grenez<sup>1</sup>, A. Kacha<sup>1</sup>, J. Schoentgen<sup>1</sup>*

<sup>1</sup>Université Libre de Bruxelles, Belgium; <sup>2</sup>Sint-Jan General Hospital, Belgium

PAGE 2226  
ThuSe4.P4-10

**Combining Neural Network and Rule-Based Systems for Dysarthria Diagnosis**

*James Carmichael, Vincent Wan, Phil Green, University of Sheffield, UK*

PAGE 2230  
ThuSe4.P4-11

**Speech as a Means of Monitoring Cognitive Function of Elderly Speakers**

*Shona D'Arcy<sup>1</sup>, Viliam Rapcan<sup>1</sup>, Nils Penard<sup>2</sup>, Margaret E. Morris<sup>3</sup>, Ian H. Robertson<sup>2</sup>, Richard B. Reilly<sup>2</sup>*

<sup>1</sup>University College Dublin, Ireland; <sup>2</sup>Trinity College Dublin, Ireland; <sup>3</sup>Intel Corporation, USA



---

*ThuSe4.P4 continued ...*

PAGE 2234  
ThuSe4.P4-12

**Integration of Metamodel and Acoustic Model for Speech Recognition**

*Hironori Matsumasa<sup>1</sup>, Tetsuya Takiguchi<sup>1</sup>, Yasuo Arika<sup>1</sup>, Ichao Li<sup>2</sup>,  
Toshitaka Nakabayashi<sup>1</sup>*

<sup>1</sup>*Kobe University, Japan;* <sup>2</sup>*Otemon Gakuin University, Japan*

PAGE 2238  
ThuSe4.P4-13

**Frequency Compression/Transposition of Fricative Consonants for the Hearing Impaired with High-Frequency Dead Regions**

*Francisco J. Fraga<sup>1</sup>, Leticia P. Costa S. Prates<sup>2</sup>, Maria Cecilia M. Iorio<sup>2</sup>*

<sup>1</sup>*UFABC, Brazil;* <sup>2</sup>*UNIFESP, Brazil*

---

---

## FriSe2.01: Automatic Speech Recognition: Features II

Great Hall, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Michael Wagner

---

PAGE 2242  
FriSe2.01-1  
10:00 - 10:20

### **On the Combination of Auditory and Modulation Frequency Channels for ASR Applications**

*Fabio Valente, Hynek Hermansky, IDIAP Research Institute, Switzerland*

PAGE 2246  
FriSe2.01-2  
10:20 - 10:40

### **Tandem Processing of Fepstrum Features**

*Vivek Tyagi, IBM India Research Lab, India*

PAGE 2250  
FriSe2.01-3  
10:40 - 11:00

### **Data-Driven Clustered Hierarchical Tandem System for LVCSR**

*Shuo-Yiin Chang, Lin-shan Lee, National Taiwan University, Taiwan*

PAGE 2254  
FriSe2.01-4  
11:00 - 11:20

### **Linear Discriminant Feature Extraction Using Weighted Classification Confusion Information**

*Hung-Shin Lee, Berlin Chen, National Taiwan Normal University, Taiwan*

PAGE 2258  
FriSe2.01-5  
11:20 - 11:40

### **Use of Spectral Centre of Gravity for Generating Speaker Invariant Features for Automatic Speech Recognition**

*D.R. Sanand, V. Balaji, Sandhya Rani R., S. Umesh, IIT Kanpur, India*

PAGE 2262  
FriSe2.01-6  
11:40 - 12:00

### **Short- and Long-Term Dynamic Features for Robust Speech Recognition**

*Takashi Fukuda, Osamu Ichikawa, Masafumi Nishimura, IBM Japan Ltd., Japan*

---

---

## FriSe2.O2: Speech Synthesis: Prosody and Emotion II

Plaza 1, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Masami Akamine

---

PAGE 2266  
FriSe2.O2-1  
10:00 - 10:20

**Duration Refinement by Jointly Optimizing State and Longer Unit Likelihood**  
*Boyang Gao, Yao Qian, Zhizheng Wu, Frank K. Soong, Microsoft Research Asia, China*

PAGE 2270  
FriSe2.O2-2  
10:20 - 10:40

**T-Tilt: A Modified Tilt Model for F0 Analysis and Synthesis in Tonal Languages**  
*Ausdang Thangthai, Nattanun Thatphithakkul, Chai Wutiwiwatchai, Anocha Rugchatjaroen, Sittipong Saychum, NECTEC, Thailand*

PAGE 2274  
FriSe2.O2-3  
10:40 - 11:00

**Multilevel Parametric-Base F0 Model for Speech Synthesis**  
*Javier Latorre, Masami Akamine, Toshiba Corporate R&D Center, Japan*

PAGE 2278  
FriSe2.O2-4  
11:00 - 11:20

**On the Generation of Synthetic Disfluent Speech: Local Prosodic Modifications Caused by the Insertion of Editing Terms**  
*Jordi Adell<sup>1</sup>, Antonio Bonafonte<sup>1</sup>, David Escudero-Mancebo<sup>2</sup>*  
*<sup>1</sup>Universitat Politècnica de Catalunya, Spain; <sup>2</sup>Universidad de Valladolid, Spain*

PAGE 2282  
FriSe2.O2-5  
11:20 - 11:40

**A Comparison of Voice Conversion Methods for Transforming Voice Quality in Emotional Speech Synthesis**  
*Oytun Türk, Marc Schröder, DFKI GmbH, Germany*

PAGE 2286  
FriSe2.O2-6  
11:40 - 12:00

**Tree Grammars as Models of Prosodic Structure**  
*Joseph Tepperman, Shrikanth S. Narayanan, University of Southern California, USA*

---

---

## FriSe2.03: Human Speech Production

Plaza 2, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Janet Fletcher

---

PAGE 2290  
FriSe2.03-1  
10:00 - 10:20

**Relation Between Geometry and Kinematics of Articulatory Trajectory Associated with Emotional Speech Production**  
*Sungbok Lee, Tsuneo Kato, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 2294  
FriSe2.03-2  
10:20 - 10:40

**Intrinsic Consonantal F0 Perturbation in 3-Way VOT Contrast and Its Implications for Aspiration-Conditioned Tonal Split: Evidence from Vietnamese**  
*Michael J. Carne, Australian National University, Australia*

PAGE 2298  
FriSe2.03-3  
10:40 - 11:00

**A Model Based Investigation of Activation Patterns of the Tongue Muscles for Vowel Production**  
*Qiang Fang, Satoru Fujita, Xugang Lu, Jianwu Dang, JAIST, Japan*

PAGE 2302  
FriSe2.03-4  
11:00 - 11:20

**Interrelationship Between Vocal Effort and Vocal Tract Acoustics: A Pilot Study**  
*Maëva Garnier<sup>1</sup>, Joe Wolfe<sup>1</sup>, Nathalie Henrich<sup>2</sup>, John Smith<sup>1</sup>*  
<sup>1</sup>University of New South Wales, Australia; <sup>2</sup>GIPSA, France

PAGE 2306  
FriSe2.03-5  
11:20 - 11:40

**Predicting Tongue Shapes from a Few Landmark Locations**  
*Chao Qin<sup>1</sup>, Miguel Á. Carreira-Perpiñán<sup>1</sup>, Korin Richmond<sup>2</sup>, Alan Wrench<sup>3</sup>, Steve Renals<sup>2</sup>*  
<sup>1</sup>University of California at Merced, USA; <sup>2</sup>University of Edinburgh, UK; <sup>3</sup>Queen Margaret University, UK

---

---

## FriSe2.04: Special Session: LIPS 2008 — Visual Speech Synthesis Challenge

Plaza 3&4, Time 10:00 – 12:00, Friday 26th September 2008

Chair: Sascha Fagel and Barry-John Theobald

---

PAGE 2310  
FriSe2.04-1

### **LIPS2008: Visual Speech Synthesis Challenge**

*Barry-John Theobald*<sup>1</sup>, *Sascha Fagel*<sup>2</sup>, *Gérard Bailly*<sup>3</sup>, *Frédéric Elisei*<sup>3</sup>

<sup>1</sup>University of East Anglia, UK; <sup>2</sup>Technische Universität Berlin, Germany; <sup>3</sup>GIPSA, France

PAGE 2314  
FriSe2.04-2

### **Speech-Driven Lip Motion Generation with a Trajectory HMM**

*Gregor Hofer*, *Junichi Yamagishi*, *Hiroshi Shimodaira*, University of Edinburgh, UK

PAGE 2318  
FriSe2.04-3

### **A Trainable Trajectory Formation Model TD-HMM Parameterized for the LIPS 2008 Challenge**

*Gérard Bailly*<sup>1</sup>, *Oxana Govokhina*<sup>1</sup>, *Gaspard Breton*<sup>2</sup>, *Frédéric Elisei*<sup>1</sup>, *Christophe Savariaux*<sup>1</sup>

<sup>1</sup>GIPSA, France; <sup>2</sup>Orange Labs, France

PAGE 2322  
FriSe2.04-4

### **Comparing Text-Driven and Speech-Driven Visual Speech Synthesizers**

*Barry-John Theobald*<sup>1</sup>, *Gavin Cawley*<sup>1</sup>, *Andrew Bangham*<sup>1</sup>, *Iain Matthews*<sup>2</sup>, *Nicholas Wilkinson*<sup>1</sup>

<sup>1</sup>University of East Anglia, UK; <sup>2</sup>Weta Digital Limited, New Zealand

PAGE 2323  
FriSe2.04-5

### **Automatic Lip Synchronization by Speech Signal Analysis**

*Goranka Zoric*, *Aleksandra Cerekovic*, *Igor S. Pandzic*, University of Zagreb, Croatia

PAGE 2324  
FriSe2.04-6

### **MASSY Speaks English: Adaptation and Evaluation of a Talking Head**

*Sascha Fagel*, Technische Universität Berlin, Germany

PAGE 2325  
FriSe2.04-7

### **From 3-D Speaker Cloning to Text-to-Audiovisual-Speech**

*Sascha Fagel*<sup>1</sup>, *Frédéric Elisei*<sup>2</sup>, *Gérard Bailly*<sup>2</sup>

<sup>1</sup>Technische Universität Berlin, Germany; <sup>2</sup>GIPSA, France

---

*FriSe2.O4 continued ...*

PAGE 2326  
FriSe2.O4-8

**A Development of Czech Talking Head**

*Zdeněk Krňoul, Miloš Železný, University of West Bohemia, Czech Republic*

PAGE 2330  
FriSe2.O4-9

**Realistic Facial Animation System for Interactive Services**

*Kang Liu, Joern Ostermann, Leibniz Universität Hannover, Germany*

PAGE 2334  
FriSe2.O4-10

**Speech-Driven 3D Facial Animation for Mobile Entertainment**

*Juan Yan, Xiang Xie, Hao Hu, Beijing Institute of Technology, China*

PAGE 2338  
FriSe2.O4-11

**A Real-Time Text to Audio-Visual Speech Synthesis System**

*Lijuan Wang<sup>1</sup>, Xiaojun Qian<sup>2</sup>, Lei Ma<sup>1</sup>, Yao Qian<sup>1</sup>, Yining Chen<sup>1</sup>, Frank K. Soong<sup>1</sup>*

*<sup>1</sup>Microsoft Research Asia, China; <sup>2</sup>Fudan University, China*

---

---

## FriSe2.05: Spoken Language Translation Systems

Plaza 5, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Yuqing Gao

---

PAGE 2342  
FriSe2.05-1  
10:00 - 10:20

**ASR Word Lattice Translation with Exhaustive Reordering is Possible**  
*Evgeny Matusov, Björn Hoffmeister, Hermann Ney, RWTH Aachen University, Germany*

PAGE 2346  
FriSe2.05-2  
10:20 - 10:40

**Development of SRI's Translation Systems for Broadcast News and Broadcast Conversations**  
*Jing Zheng, Wen Wang, Necip Fazil Ayan, SRI International, USA*

PAGE 2350  
FriSe2.05-3  
10:40 - 11:00

**Machine Translation in Continuous Space**  
*Ruhi Sarikaya, Yonggang Deng, Mohamed Afify, Brian Kingsbury, Yuqing Gao, IBM T.J. Watson Research Center, USA*

PAGE 2354  
FriSe2.05-4  
11:00 - 11:20

**Discovering Phrases in Machine Translation by Simulated Annealing**  
*Caroline Lavecchia, David Langlois, Kamel Smâili, LORIA, France*

PAGE 2358  
FriSe2.05-5  
11:20 - 11:40

**Towards Domain Independence in Machine Aided Human Translation**  
*Aarthi Reddy, Richard C. Rose, McGill University, Canada*

PAGE 2362  
FriSe2.05-6  
11:40 - 12:00

**Class-Based Statistical Machine Translation for Field Maintainable Speech-to-Speech Translation**  
*Ian R. Lane, Alex Waibel, Mobile Technologies LLC, USA*

---

---

## FriSe2.P1 : Automatic Speech Recognition: Acoustic Models III

Mezzanine Level Area A1, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Lori Lamel

---

PAGE 2366  
FriSe2.P1-1

### **Nonnative Speech Recognition Based on State-Candidate Bilingual Model Modification**

*Qingqing Zhang, Ta Li, Jielin Pan, Yonghong Yan, Chinese Academy of Sciences, China*

PAGE 2370  
FriSe2.P1-2

### **Prosodic and Spectral Features Within Segment-Based Acoustic Modeling**

*Björn Schuller, Xiaohua Zhang, Gerhard Rigoll, Technische Universität München, Germany*

PAGE 2374  
FriSe2.P1-3

### **Unsupervised versus Supervised Training of Acoustic Models**

*Jeff Ma, Richard Schwartz, BBN Technologies, USA*

PAGE 2378  
FriSe2.P1-4

### **A Comparison of Broad Phonetic and Acoustic Units for Noise Robust Segment-Based Phonetic Recognition**

*Tara N. Sainath, Victor Zue, MIT, USA*

PAGE 2382  
FriSe2.P1-5

### **Aggregated Cross-Validation and Its Efficient Application to Gaussian Mixture Optimization**

*Takahiro Shinozaki<sup>1</sup>, Sadaoki Furui<sup>1</sup>, Tatsuya Kawahara<sup>2</sup>*

*<sup>1</sup>Tokyo Institute of Technology, Japan; <sup>2</sup>Kyoto University, Japan*

PAGE 2386  
FriSe2.P1-6

### **A Minimum Classification Error Based Distance Measure for Template Based Speech Recognition**

*Mike Matton, Dirk Van Compernelle, Ronald Cools, Katholieke Universiteit Leuven, Belgium*



---

*FriSe2.P1 continued ...*

PAGE 2390  
FriSe2.P1-7

**A Penalized Logistic Regression Approach to Detection Based Phone Classification**

*Sabato Marco Siniscalchi<sup>1</sup>, Torbjørn Svendsen<sup>1</sup>, Chin-Hui Lee<sup>2</sup>*

*<sup>1</sup>NTNU, Norway; <sup>2</sup>Georgia Institute of Technology, USA*

PAGE 2394  
FriSe2.P1-8

**Incorporating Acoustical Modelling of Phone Transitions in an Hybrid ANN/HMM Speech Recognizer**

*Alberto Abad, João Neto, INESC-ID/IST, Portugal*

PAGE 2398  
FriSe2.P1-9

**Flexible Discriminative Training Based on Equal Error Group Scores Obtained from an Error-Indexed Forward-Backward Algorithm**

*Erik McDermott, Atsushi Nakamura, NTT Corporation, Japan*

PAGE 2402  
FriSe2.P1-10

**Pitch Adaptive Features for LVCSR**

*Giulia Garau, Steve Renals, University of Edinburgh, UK*

PAGE 2406  
FriSe2.P1-11

**Using Syllable Nuclei Locations to Improve Automatic Speech Recognition in the Presence of Burst Noise**

*Chris D. Bartels, Jeff A. Bilmes, University of Washington, USA*

PAGE 2410  
FriSe2.P1-12

**Effects of Allophones on the Performance of Korean Speech Recognition**

*Hyejin Hong, Sunhee Kim, Minhwa Chung, Seoul National University, Korea*

PAGE 2414  
FriSe2.P1-13

**Combining Evidence from a Generative and a Discriminative Model in Phoneme Recognition**

*Joel Pinto, Hynek Hermansky, IDIAP Research Institute, Switzerland*

PAGE 2418  
FriSe2.P1-14

**Fragmented Context-Dependent Syllable Acoustic Models**

*K. Thambiratnam, Frank Seide, Microsoft Research Asia, China*

---

*FriSe2.P1 continued ...*

PAGE 2422  
FriSe2.P1-15

**Speech Recognition Using Non-Linear Trajectories in a Formant-Based  
Articulatory Layer of a Multiple-Level Segmental HMM**

*Hongwei Hu, Martin J. Russell, University of Birmingham, UK*

PAGE 2426  
FriSe2.P1-16

**Recent Improvements of the RWTH GALE Mandarin LVCSR System**

*Ch. Plahl<sup>1</sup>, Björn Hoffmeister<sup>1</sup>, M.-Y. Hwang<sup>2</sup>, D. Lu<sup>3</sup>, Georg Heigold<sup>1</sup>, Jonas Löff<sup>1</sup>,  
Ralf Schlüter<sup>1</sup>, Hermann Ney<sup>1</sup>*

*<sup>1</sup>RWTH Aachen University, Germany; <sup>2</sup>Microsoft Research, USA; <sup>3</sup>Southwest Forestry  
University, China*

---

---

## FriSe2.P2: Spoken Language: Parsing and Summarisation

Mezzanine Level Area A2, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Pietro Laface

---

PAGE 2430  
FriSe2.P2-1

### **Intonational Phrases for Speech Summarization**

*Sameer R. Maskey<sup>1</sup>, Andrew Rosenberg<sup>2</sup>, Julia Hirschberg<sup>2</sup>*

<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>Columbia University, USA

PAGE 2434  
FriSe2.P2-2

### **Packing the Meeting Summarization Knapsack**

*Korbinian Riedhammer<sup>1</sup>, Dan Gillick<sup>2</sup>, Benoit Favre<sup>3</sup>, Dilek Hakkani-Tür<sup>3</sup>*

<sup>1</sup>Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; <sup>2</sup>University of California at Berkeley, USA; <sup>3</sup>ICSI, USA

PAGE 2438  
FriSe2.P2-3

### **Class Lecture Summarization Taking into Account Consecutiveness of Important Sentences**

*Yasuhisa Fujii<sup>1</sup>, Kazumasa Yamamoto<sup>1</sup>, Norihide Kitaoka<sup>2</sup>, Seiichi Nakagawa<sup>1</sup>*

<sup>1</sup>Toyohashi University of Technology, Japan; <sup>2</sup>Nagoya University, Japan

PAGE 2442  
FriSe2.P2-4

### **Using Latent Dirichlet Allocation to Incorporate Domain Knowledge for Topic Transition Detection**

*Xiaodan Zhu, Xuming He, Cosmin Munteanu, Gerald Penn, University of Toronto, Canada*

PAGE 2446  
FriSe2.P2-5

### **Weakly Supervised Training for Parsing Mandarin Broadcast Transcripts**

*Wen Wang, SRI International, USA*

PAGE 2450  
FriSe2.P2-6

### **Parsing with Subdomain Instance Weighting from Raw Corpora**

*Barbara Plank<sup>1</sup>, Khalil Sima'an<sup>2</sup>*

<sup>1</sup>University of Groningen, The Netherlands; <sup>2</sup>University of Amsterdam, The Netherlands

---

*FriSe2.P2 continued ...*

PAGE 2454  
FriSe2.P2-7

**Dependency Parsing of Japanese Spoken Monologue Based on Clause-Starts Detection**

*Tomohiro Ohno*<sup>1</sup>, *Shigeki Matsubara*<sup>1</sup>, *Hideki Kashioka*<sup>2</sup>, *Yasuyoshi Inagaki*<sup>3</sup>

<sup>1</sup>*Nagoya University, Japan*; <sup>2</sup>*NICT, Japan*; <sup>3</sup>*Toyohashi University of Technology, Japan*

PAGE 2458  
FriSe2.P2-8

**Online Unsupervised Pattern Discovery in Speech Using Parallelization**

*Mrugesh R. Gajjar*, *R. Govindarajan*, *T.V. Sreenivas*, *Indian Institute of Science, India*

---

---

## FriSe2.P3 : Multimodal Interfaces

Mezzanine Level Area B3, Time 10:00 - 12:00, Friday 26th September 2008

Chair: David House

---

PAGE 2462  
FriSe2.P3-1

### **A Comparison of Input Entry Rates in a Multimodal Mobile Application**

*Aleksi Melto, Markku Turunen, Jaakko Hakulinen, Anssi Kainulainen, Tomi Heimonen, University of Tampere, Finland*

PAGE 2466  
FriSe2.P3-2

### **Physically Embodied Conversational Agents as Health and Fitness Companions**

*Markku Turunen<sup>1</sup>, Jaakko Hakulinen<sup>1</sup>, Cameron Smith<sup>2</sup>, Daniel Charlton<sup>2</sup>, Li Zhang<sup>2</sup>, Marc Cavazza<sup>2</sup>*

<sup>1</sup>University of Tampere, Finland; <sup>2</sup>University of Teesside, UK

PAGE 2470  
FriSe2.P3-3

### **User Perception of Multi-Modal Interfaces for Mobile Applications**

*Florian Metzke<sup>1</sup>, Roman Englert<sup>2</sup>, Udo Bub<sup>3</sup>, Ingmar Kliche<sup>4</sup>, Thomas Scheerbarth<sup>4</sup>*

<sup>1</sup>Technische Universität Berlin, Germany; <sup>2</sup>Ben-Gurion University of the Negev, Israel;

<sup>3</sup>Deutsche Telekom Laboratories, Germany; <sup>4</sup>T-Systems Enterprise Services GmbH, Germany

PAGE 2474  
FriSe2.P3-4

### **Design and Formulation for Speech Interface Based on Flexible Shortcuts**

*Tepei Nakano<sup>1</sup>, Tomoyuki Kumai<sup>1</sup>, Tetsunori Kobayashi<sup>1</sup>, Yasushi Ishikawa<sup>2</sup>*

<sup>1</sup>Waseda University, Japan; <sup>2</sup>Mitsubishi Electric Corp., Japan

PAGE 2478  
FriSe2.P3-5

### **Exploring Classification Techniques in Speech Based Cognitive Load Monitoring**

*Bo Yin<sup>1</sup>, Natalie Ruiz<sup>2</sup>, Fang Chen<sup>2</sup>, Eliathamby Ambikairajah<sup>1</sup>*

<sup>1</sup>University of New South Wales, Australia; <sup>2</sup>NICTA, Australia

PAGE 2482  
FriSe2.P3-6

### **Finding Two-Level Interpersonal Context: Proximity and Conversation Detection from Personal Audio Feature Data**

*Masayuki Okamoto, Naoki Iketani, Keisuke Nishimura, Masaaki Kikuchi, Kenta Cho, Masanori Hattori, Sougo Tsuboi, Toshiba Corporate R&D Center, Japan*

---

*FriSe2.P3 continued ...*

PAGE 2486  
FriSe2.P3-7

**From Domain Specification to Virtual Humans: An Integrated Approach to Authoring Tactical Questioning Characters**

*Sudeep Gandhe, David DeVault, Antonio Roque, Bilyana Martinovski, Ron Artstein, Anton Leuski, Jillian Gerten, David Traum, University of Southern California, USA*

PAGE 2490  
FriSe2.P3-8

**Designing a Massively Multiplayer Online Role-Playing Game Around Text-to-Speech**

*Mike Rozak,  $\mu$ Xac, Australia*

---

---

## FriSe2.P4: Speech, Music, Audio Segmentation and Classification

Mezzanine Level Area B4, Time 10:00 - 12:00, Friday 26th September 2008

Chair: Haizhou Li

---

PAGE 2494  
FriSe2.P4-1

### **Robust Speaker Change Detection Using Kernel-Gaussian Model**

*Jie Gao, Xiang Zhang, Qingwei Zhao, Yonghong Yan, Chinese Academy of Sciences, China*

PAGE 2498  
FriSe2.P4-2

### **A Comparative Study in Automatic Recognition of Broadcast Audio**

*Stavros Ntalampiras, Nikos Fakotakis, University of Patras, Greece*

PAGE 2502  
FriSe2.P4-3

### **Joint Time-Frequency Segmentation for Transient Decomposition**

*Charturong Tantibundhit<sup>1</sup>, Gernot Kubin<sup>2</sup>*

*<sup>1</sup>Thammasat University, Thailand; <sup>2</sup>Graz University of Technology, Austria*

PAGE 2506  
FriSe2.P4-4

### **Language and Genre Detection in Audio Content Analysis**

*Vikramjit Mitra, Daniel Garcia-Romero, Carol Y. Espy-Wilson, University of Maryland, USA*

PAGE 2510  
FriSe2.P4-5

### **An Entropy Based Feature for Whisper-Island Detection Within Audio Streams**

*Chi Zhang, John H.L. Hansen, University of Texas at Dallas, USA*

PAGE 2514  
FriSe2.P4-6

### **Two Step Speaker Segmentation Method Using Bayesian Information Criterion and Adapted Gaussian Mixtures Models**

*Matej Grašič, Marko Kos, Andrej Žgank, Zdravko Kačič, University of Maribor, Slovenia*

PAGE 2518  
FriSe2.P4-7

### **Domain-Specific Classification Methods for Disfluency Detection**

*Sebastian Germesin, Tilman Becker, Peter Poller, DFKI GmbH, Germany*

PAGE 2522  
FriSe2.P4-8

### **Multi-Speaker Meeting Audio Segmentation**

*Tin Lay Nwe, Minghui Dong, Swe Zin Kalayar Khine, Haizhou Li, Institute for Infocomm Research, Singapore*

---

*FriSe2.P4 continued ...*

PAGE 2526  
FriSe2.P4-9

**Rhythm Based Music Segmentation and Octave Scale Cepstral Features for Sung Language Recognition**

*Namunu C. Maddage, Haizhou Li, Institute for Infocomm Research, Singapore*

PAGE 2530  
FriSe2.P4-10

**Robust Voiced/Unvoiced Speech Classification Using Empirical Mode Decomposition and Periodic Correlation Model**

*Md. Khademul Islam Molla, Keikichi Hirose, Nobuaki Minematsu, University of Tokyo, Japan*

PAGE 2534  
FriSe2.P4-11

**A Combination of Data Mining Method with Decision Trees Building for Speech/Music Discrimination**

*Qiong Wu<sup>1</sup>, Qin Yan<sup>2</sup>, Jun Wang<sup>1</sup>, Jun Hong<sup>1</sup>*

*<sup>1</sup>Chinese Academy of Sciences, China; <sup>2</sup>Hohai University, China*

PAGE 2538  
FriSe2.P4-12

**Advertisement Detection in French Broadcast News Using Acoustic Repetition and Gaussian Mixture Models**

*Vishwa Gupta, Gilles Boulianne, Patrick Kenny, Pierre Dumouchel, CRIM, Canada*

PAGE 2542  
FriSe2.P4-13

**A Hybrid SVM/MCE Training Approach for Vector Space Topic Identification of Spoken Audio Recordings**

*Timothy J. Hazen, Fred Richardson, MIT, USA*

PAGE 2546  
FriSe2.P4-14

**Training Audio Events Detectors with a Sound Effects Corpus**

*Isabel Trancoso<sup>1</sup>, José Portêlo<sup>1</sup>, Miguel Bugalho<sup>1</sup>, João Neto<sup>1</sup>, António Serralheiro<sup>2</sup>*

*<sup>1</sup>INESC-ID/IST, Portugal; <sup>2</sup>INESC-ID/Academia Militar, Portugal*



---

---

## FriSe3.O1 : Automatic Speech Recognition: New Paradigms

Great Hall, Time 13:30 - 15:30, Friday 26th September 2008

Chair: Alex Acero

---

PAGE 2550  
FriSe3.O1-1  
13:30 - 13:50

### **Longitudinal Study of ASR Performance on Ageing Voices**

*Ravichander Vipperla, Steve Renals, Joe Frankel, University of Edinburgh, UK*

PAGE 2554  
FriSe3.O1-2  
13:50 - 14:10

### **HAC-Models: A Novel Approach to Continuous Speech Recognition**

*Hugo Van hamme, Katholieke Universiteit Leuven, Belgium*

PAGE 2558  
FriSe3.O1-3  
14:10 - 14:30

### **Investigations into Phonological Attribute Classifier Representations for CRF Phone Recognition**

*Prateeti Mohapatra, Eric Fosler-Lussier, Ohio State University, USA*

PAGE 2562  
FriSe3.O1-4  
14:30 - 14:50

### **Applications of Virtual-Evidence Based Speech Recognizer Training**

*Amarnag Subramanya, Jeff A. Bilmes, University of Washington, USA*

PAGE 2566  
FriSe3.O1-5  
14:50 - 15:10

### **Spoken Digit Recognition Using a Hierarchical Temporal Memory**

*Joost van Doremalen, Lou Boves, Radboud Universiteit Nijmegen, The Netherlands*

PAGE 2570  
FriSe3.O1-6  
15:10 - 15:30

### **A Computational Model of Language Acquisition: Focus on Word Discovery**

*Louis ten Bosch<sup>1</sup>, Hugo Van hamme<sup>2</sup>, Lou Boves<sup>1</sup>*

*<sup>1</sup>Radboud Universiteit Nijmegen, The Netherlands; <sup>2</sup>Katholieke Universiteit Leuven, Belgium*

---

---

## FriSe3.O2: Speech and Acoustic Activity Detection

Plaza 1, Time 13:30 - 15:30, Friday 26th September 2008

Chair: Jean-Pierre Martens

---

PAGE 2574  
FriSe3.O2-1  
13:30 - 13:50

### **Voice Activity Detection Using Modified Wigner-Ville Distribution**

*Lakshmish Kaushik, Douglas O'Shaughnessy, Université du Québec, Canada*

PAGE 2578  
FriSe3.O2-2  
13:50 - 14:10

### **Energy and Entropy Based Switching Algorithm for Speech Endpoint Detection in Varying SNR Conditions**

*Krishna Chaitanya, Rohit Sinha, IIT Guwahati, India*

PAGE 2582  
FriSe3.O2-3  
14:10 - 14:30

### **Detection of Speech Embedded in Real Acoustic Background Based on Amplitude Modulation Spectrogram Features**

*Jörn Anemüller, Denny Schmidt, Jörg-Hendrik Bach, Carl von Ossietzky Universität Oldenburg, Germany*

PAGE 2586  
FriSe3.O2-4  
14:30 - 14:50

### **Voice Activity Detection Algorithms Using Subband Power Distance Feature for Noisy Environments**

*Tuan Van Pham, Michael Stadtschnitzer, Franz Pernkopf, Gernot Kubin, Graz University of Technology, Austria*

PAGE 2590  
FriSe3.O2-5  
14:50 - 15:10

### **Speech-Overlapped Acoustic Event Detection for Automotive Applications**

*Christian Müller<sup>1</sup>, Joan-Isaac Biel<sup>1</sup>, Edward Kim<sup>2</sup>, Daniel Rosario<sup>2</sup>*

*<sup>1</sup>ICSI, USA; <sup>2</sup>Volkswagen of America, USA*

PAGE 2594  
FriSe3.O2-6  
15:10 - 15:30

### **Detection of Acoustic Events in Interactive Seminar Data with Temporal Overlaps**

*Andrey Temko, Climent Nadeu, Universitat Politècnica de Catalunya, Spain*

---

---

## FriSe3.03: Speech Analysis and Processing

Plaza 2, Time 13:30 - 15:30, Friday 26th September 2008

Chair: Catherine I. Watson

---

PAGE 2598  
FriSe3.03-1  
13:30 - 13:50

### **Robust Signal-to-Noise Ratio Estimation Based on Waveform Amplitude Distribution Analysis**

*Chanwoo Kim, Richard M. Stern, Carnegie Mellon University, USA*

PAGE 2602  
FriSe3.03-2  
13:50 - 14:10

### **Speech Analysis Using Instantaneous Frequency Deviation**

*Anthony P. Stark, Kuldip K. Paliwal, Griffith University, Australia*

PAGE 2606  
FriSe3.03-3  
14:10 - 14:30

### **Auditory-Based Formant Estimation in Noise Using a Probabilistic Framework**

*Claudius Gläser, Martin Heckmann, Frank Joublin, Christian Goerick, Honda Research Institute Europe GmbH, Germany*

PAGE 2610  
FriSe3.03-4  
14:30 - 14:50

### **Efficient Representation of Throat Microphone Speech**

*Sri Rama Murty K.<sup>1</sup>, Saurav Khurana<sup>2</sup>, Yogendra Umesh Itankar<sup>2</sup>, M.R. Kesheorey<sup>3</sup>, B. Yegnanarayana<sup>2</sup>*

*<sup>1</sup>IIT Madras, India; <sup>2</sup>IIT Hyderabad, India; <sup>3</sup>Center for Artificial Intelligence & Robotics, India*

PAGE 2614  
FriSe3.03-5  
14:50 - 15:10

### **Acoustic-Phonetic Approach for Automatic Evaluation of Spoken Grammar**

*Om D. Deshmukh, Ashish Verma, IBM India Research Lab, India*

PAGE 2618  
FriSe3.03-6  
15:10 - 15:30

### **On Estimation of a Speaker's Confusion Matrix from Sparse Data**

*Stephen Cox, University of East Anglia, UK*

---

---

## FriSe3.04: Special Session: Talking Heads and Pronunciation Training

Plaza 3&4, Time 13:30 – 15:30, Friday 26th September 2008

Chair: Gérard Bailly

---

PAGE 2622  
FriSe3.04-1

**Talking Heads and Pronunciation Training: A Review**  
*Valerie Hazan, University College London, UK*

PAGE 2623  
FriSe3.04-2

**Pronunciation Training: The Role of Eye and Ear**  
*Dominic W. Massaro<sup>1</sup>, Stephanie Bigler<sup>1</sup>, Trevor Chen<sup>1</sup>, Marcus Perlman<sup>1</sup>, Slim Ouni<sup>2</sup>*  
<sup>1</sup>University of California at Santa Cruz, USA; <sup>2</sup>LORIA, France

PAGE 2627  
FriSe3.04-3

**Can Visualization of Internal Articulators Support Speech Perception?**  
*Preben Wik, Olov Engwall, KTH, Sweden*

PAGE 2631  
FriSe3.04-4

**Can Audio-Visual Instructions Help Learners Improve Their Articulation? — An Ultrasound Study of Short Term Changes**  
*Olov Engwall, KTH, Sweden*

PAGE 2635  
FriSe3.04-5

**Can You “Read Tongue Movements”?**  
*Pierre Badin, Yuliya Tarabalka, Frédéric Elisei, Gérard Bailly, GIPSA, France*

PAGE 2639  
FriSe3.04-6

**Two- and Three-Dimensional Visual Articulatory Models for Pronunciation Training and for Treatment of Speech Disorders**  
*Bernd J. Kröger<sup>1</sup>, Verena Graf-Borttscheller<sup>1</sup>, Anja Lowit<sup>2</sup>*  
<sup>1</sup>RWTH Aachen University, Germany; <sup>2</sup>University of Strathclyde, UK

PAGE 2643  
FriSe3.04-7

**A 3-D Virtual Head as a Tool for Speech Therapy for Children**  
*Sascha Fagel, Katja Madany, Technische Universität Berlin, Germany*

PAGE 2647  
FriSe3.04-8

**AnTon: An Animatronic Model of a Human Tongue and Vocal Tract**  
*Robin Hofe, Roger K. Moore, University of Sheffield, UK*

---

*FriSe3.O4 continued ...*

PAGE 2651  
FriSe3.O4-9

**Physical Models of the Human Vocal Tract with Gel-Type Material**

*Takayuki Arai, Sophia University, Japan*

PAGE 2655  
FriSe3.O4-10

**Mispronunciation Detection for Mandarin Chinese**

*Chao Huang, Feng Zhang, Frank K. Soong, Min Chu, Microsoft Research Asia, China*

---

---

## FriSe3.05: Multimodal Speech Processing

Plaza 5, Time 13:30 - 15:30, Friday 26th September 2008

Chair: Takaaki Kuratate

---

PAGE 2659  
FriSe3.05-1  
13:30 - 13:50

### **Efficient Handwriting Correction of Speech Recognition Errors with Template Constrained Posterior (TCP)**

*Lijuan Wang<sup>1</sup>, Tao Hu<sup>2</sup>, Peng Liu<sup>1</sup>, Frank K. Soong<sup>1</sup>*

*<sup>1</sup>Microsoft Research Asia, China; <sup>2</sup>Shanghai Jiao Tong University, China*

PAGE 2663  
FriSe3.05-2  
13:50 - 14:10

### **Bi-Gaussian Score Equalization in an Audio-Visual SVM-Based Person Verification System**

*Pascual Ejarque, Javier Hernando, Universitat Politècnica de Catalunya, Spain*

PAGE 2667  
FriSe3.05-3  
14:10 - 14:30

### **Speech Recognition for Vocalized and Subvocal Modes of Production Using Surface EMG Signals from the Neck and Face**

*Geoffrey S. Meltzner<sup>1</sup>, Jason Sroka<sup>1</sup>, James T. Heaton<sup>2</sup>, L. Donald Gilmore<sup>3</sup>, Glen Colby<sup>1</sup>, Serge Roy<sup>3</sup>, Nancy Chen<sup>1</sup>, Carlo J. De Luca<sup>3</sup>*

*<sup>1</sup>BAE Systems, USA; <sup>2</sup>Massachusetts General Hospital, USA; <sup>3</sup>Altec Inc., USA*

PAGE 2671  
FriSe3.05-4  
14:30 - 14:50

### **Distinctive Feature Fusion for Recognition of Australian English Consonants**

*Trent W. Lewis, David M.W. Powers, Flinders University, Australia*

PAGE 2675  
FriSe3.05-5  
14:50 - 15:10

### **Time-Lag Adaptation for Semi-Synchronous Speech and Pen Input**

*Yasushi Watanabe, Koichi Shinoda, Sadaoki Furui, Tokyo Institute of Technology, Japan*

PAGE 2679  
FriSe3.05-6  
15:10 - 15:30

### **Continuous Pose-Invariant Lipreading**

*Patrick Lucey, Sridha Sridharan, David Dean, Queensland University of Technology, Australia*

---

---

## FriSe3.P1 : Cross-Lingual and Multilingual Automatic Speech Recognition, Speech Translation

Mezzanine Level Area A1, Time 13:30 - 15:30, Friday 26th September 2008

Chair: Alex Waibel

---

- PAGE 2683  
FriSe3.P1-1      **Czech-to-Slovak Adapted Broadcast News Transcription System**  
*Jan Nouza, Jan Silovsky, Jindrich Zdansky, Petr Cerva, Martin Kroul, Josef Chaloupka, Technical University of Liberec, Czech Republic*
- PAGE 2687  
FriSe3.P1-2      **Continuous Phone Recognition Without Target Language Training Data**  
*Dau-Cheng Lyu<sup>1</sup>, Sabato Marco Siniscalchi<sup>2</sup>, Tae-Yoon Kim<sup>3</sup>, Chin-Hui Lee<sup>3</sup>*  
*<sup>1</sup>Chang Gung University, Taiwan; <sup>2</sup>NTNU, Norway; <sup>3</sup>Georgia Institute of Technology, USA*
- PAGE 2691  
FriSe3.P1-3      **An Investigation of Acoustic Models for Multilingual Code-Switching**  
*Christopher M. White, Sanjeev Khudanpur, James K. Baker, Johns Hopkins University, USA*
- PAGE 2695  
FriSe3.P1-4      **Cross-Lingual Portability of MLP-Based Tandem Features — A Case Study for English and Hungarian**  
*László Tóth<sup>1</sup>, Joe Frankel<sup>2</sup>, Gábor Gosztolya<sup>1</sup>, Simon King<sup>2</sup>*  
*<sup>1</sup>Hungarian Academy of Sciences, Hungary; <sup>2</sup>University of Edinburgh, UK*
- PAGE 2699  
FriSe3.P1-5      **Seed Models Combination and State Level Mappings of Cross-Lingual Transfer for Rapid HMM Development: From English to Mandarin**  
*Xufang Zhao, Douglas O'Shaughnessy, Université du Québec, Canada*
- PAGE 2703  
FriSe3.P1-6      **Multi-Accent and Accent-Independent Non-Native Speech Recognition**  
*Ghazi Bouselmi, Dominique Fohr, Irina Illina, LORIA, France*
- PAGE 2707  
FriSe3.P1-7      **Cross-Lingual Sentence Extraction for Information Distillation**  
*Adish Kumar Singla, Dilek Hakkani-Tür, ICSI, USA*

---

*FriSe3.P1 continued ...*

PAGE 2711  
FriSe3.P1-8

**On the Use of a Multilingual Neural Network Front-End**

*Stefano Scanzio<sup>1</sup>, Pietro Laface<sup>1</sup>, Luciano Fissore<sup>2</sup>, Roberto Gemello<sup>2</sup>, Franco Mana<sup>2</sup>*

<sup>1</sup>Politecnico di Torino, Italy; <sup>2</sup>Loquendo, Italy

PAGE 2715  
FriSe3.P1-9

**Context-Sensitive Probabilistic Phone Mapping Model for Cross-Lingual Speech Recognition**

*Khe Chai Sim, Haizhou Li, Institute for Infocomm Research, Singapore*

PAGE 2719  
FriSe3.P1-10

**A Non-Acoustic Approach to Crosslingual Speech Recognition Performance Prediction**

*Chen Liu, Lynette Melnar, Motorola, USA*

PAGE 2723  
FriSe3.P1-11

**Factored Translation Models for Enriching Spoken Language Translation with Prosody**

*Vivek Kumar Rangarajan Sridhar<sup>1</sup>, Srinivas Bangalore<sup>2</sup>, Shrikanth S. Narayanan<sup>1</sup>*

<sup>1</sup>University of Southern California, USA; <sup>2</sup>AT&T Labs Research, USA

PAGE 2727  
FriSe3.P1-12

**Data Selection and Smoothing in an Open-Source System for the 2008 NIST Machine Translation Evaluation**

*Holger Schwenk, Yannick Estève, LIUM, France*

PAGE 2731  
FriSe3.P1-13

**Strategies for Building a Farsi-English SMT System from Limited Resources**

*Andreas Kathol, Jing Zheng, SRI International, USA*

PAGE 2735  
FriSe3.P1-14

**Stream Decoding for Simultaneous Spoken Language Translation**

*Muntsin Kolss<sup>1</sup>, Stephan Vogel<sup>2</sup>, Alex Waibel<sup>1</sup>*

<sup>1</sup>Universität Karlsruhe (TH), Germany; <sup>2</sup>Carnegie Mellon University, USA



---

*FriSe3.P1 continued ...*

PAGE 2739  
FriSe3.P1-15

**Towards Unsupervised Training of the Classifier-Based Speech Translator**

*Emil Ettelaie, Panayiotis G. Georgiou, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 2743  
FriSe3.P1-16

**Aggregating Distributed STT, MT, and Information Extraction Engines: The GALE Interoperability-Demo System**

*John F. Pitrelli<sup>1</sup>, Burn L. Lewis<sup>1</sup>, Edward A. Epstein<sup>1</sup>, Martin Franz<sup>1</sup>, Daniel Kiecza<sup>2</sup>, Jerome L. Quinn<sup>1</sup>, Ganesh Ramaswamy<sup>1</sup>, Amit Srivastava<sup>2</sup>, Paola Virga<sup>1</sup>*

<sup>1</sup>IBM T.J. Watson Research Center, USA; <sup>2</sup>BBN Technologies, USA

---

---

## FriSe3.P2 : Expression, Emotion and Personality Recognition

Mezzanine Level Area A2, Time 13:30 - 15:30, Friday 26th September 2008

Chair: Kjell Elenius

---

PAGE 2747  
FriSe3.P2-1

### **An Interval Type-2 Fuzzy Logic System to Translate Between Emotion-Related Vocabularies**

*Abe Kazemzadeh, Sungbok Lee, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 2751  
FriSe3.P2-2

### **Applying Pitch-Dependent Difference Detection and Modification to Emotional Speaker Recognition**

*Ting Huang, Yingchun Yang, Zhejiang University, China*

PAGE 2755  
FriSe3.P2-3

### **Automatic Recognition of Anger in Spontaneous Speech**

*Daniel Neiberg, Kjell Elenius, KTH, Sweden*

PAGE 2759  
FriSe3.P2-4

### **An Estimation Technique of Style Expressiveness for Emotional Speech Using Model Adaptation Based on Multiple-Regression HSMM**

*Takashi Nose, Yoichi Kato, Makoto Tachibana, Takao Kobayashi, Tokyo Institute of Technology, Japan*

PAGE 2763  
FriSe3.P2-5

### **A Vowel Based Approach for Acted Emotion Recognition**

*Fabien Ringeval, Mohamed Chetouani, ISIR, France*

PAGE 2767  
FriSe3.P2-6

### **A Composite Framework for Affective Sensing**

*Gordon McIntyre, Roland Goecke, Australian National University, Australia*

PAGE 2771  
FriSe3.P2-7

### **Towards Automatic Emotional State Categorization from Speech Signals**

*Arslan Shaukat, Ke Chen, University of Manchester, UK*

PAGE 2775  
FriSe3.P2-8

### **Speaker-Independent Emotion Recognition Based on Feature Vector Classification**

*Jeong-Sik Park<sup>1</sup>, Ji-Hwan Kim<sup>2</sup>, Sang-Min Yoon<sup>1</sup>, Yung-Hwan Oh<sup>1</sup>*

<sup>1</sup>KAIST, Korea; <sup>2</sup>Sogang University, Korea

---

---

## FriSe3.P3: Applications in Education and Learning II

Mezzanine Level Area B3, Time 13:30 – 15:30, Friday 26th September 2008

Chair: Gordon McIntyre

---

PAGE 2779  
FriSe3.P3-1

### **Estimation of Children's Reading Ability by Fusion of Automatic Pronunciation Verification and Fluency Detection**

*Matthew Black, Joseph Tepperman, Sungbok Lee, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 2783  
FriSe3.P3-2

### **Pronunciation Verification of English Letter-Sounds in Preliterate Children**

*Matthew Black, Joseph Tepperman, Abe Kazemzadeh, Sungbok Lee, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 2787  
FriSe3.P3-3

### **Improving Mispronunciation Detection and Diagnosis of Learners' Speech with Context-Sensitive Phonological Rules Based on Language Transfer**

*Alissa M. Harrison<sup>1</sup>, Wing Yiu Lau<sup>1</sup>, Helen M. Meng<sup>1</sup>, Lan Wang<sup>2</sup>  
<sup>1</sup>Chinese University of Hong Kong, China; <sup>2</sup>Chinese Academy of Sciences, China*

PAGE 2791  
FriSe3.P3-4

### **DISCO: Development and Integration of Speech Technology into Courseware for Language Learning**

*Catia Cucchiarini, Joost van Doremalen, Helmer Strik, Radboud Universiteit Nijmegen, The Netherlands*

PAGE 2795  
FriSe3.P3-5

### **Discriminative Model Combination and Language Model Selection in a Reading Tutor for Children**

*Abdurrahman Samir, Jacques Duchateau, Hugo Van hamme, Katholieke Universiteit Leuven, Belgium*

PAGE 2799  
FriSe3.P3-6

### **Usability of ASR-Based Reading Training for Dyslexics**

*Jakob Schou Pedersen, Lars Bo Larsen, Børge Lindberg, Aalborg University, Denmark*

---

*FriSe3.P3 continued ...*

PAGE 2803  
FriSe3.P3-7

**A Browsing System for Classroom Lecture Speech**

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

PAGE 2807  
FriSe3.P3-8

**Automatic Pronunciation Evaluation of Language Learners' Utterances Generated Through Shadowing**

*Dean Luo<sup>1</sup>, Naoya Shimomura<sup>1</sup>, Nobuaki Minematsu<sup>1</sup>, Yutaka Yamauchi<sup>2</sup>, Keikichi Hirose<sup>1</sup>*

*<sup>1</sup>University of Tokyo, Japan; <sup>2</sup>Tokyo International University, Japan*

PAGE 2811  
FriSe3.P3-9

**Application and Evaluation of Speech Technologies in Language Learning: Experiments with the Saybot Player**

*Sylvain Chevalier, Zhenhai Cao, Saybot Inc., China*

PAGE 2815  
FriSe3.P3-10

**Forward Optimal Modeling of Acoustic Confusions in Mandarin CALL System**

*Fengpei Ge, Fuping Pan, Changliang Liu, Bin Dong, Yonghong Yan, Chinese Academy of Sciences, China*

PAGE 2819  
FriSe3.P3-11

**Recognition of English Utterances with Grammatical and Lexical Mistakes for Dialogue-Based CALL System**

*Akinori Ito<sup>1</sup>, Ryohei Tsutsui<sup>1</sup>, Shozo Makino<sup>1</sup>, Motoyuki Suzuki<sup>2</sup>*

*<sup>1</sup>Tohoku University, Japan; <sup>2</sup>University of Tokushima, Japan*

---

---

## FriSe3.P4: Human Speech Production and Speech Perception

Mezzanine Level Area B4, Time 13:30 – 15:30, Friday 26th September 2008

Chair: Eva Hajičová

---

PAGE 2823  
FriSe3.P4-1

### **An Analysis of Vocal Tract Shaping in English Sibilant Fricatives Using Real-Time Magnetic Resonance Imaging**

*Erik Bresch, Daylen Riggs, Louis M. Goldstein, Dani Byrd, Sungbok Lee, Shrikanth S. Narayanan, University of Southern California, USA*

PAGE 2827  
FriSe3.P4-2

### **Science Workshop with Sliding Vocal-Tract Model**

*Takayuki Arai, Sophia University, Japan*

PAGE 2831  
FriSe3.P4-3

### **Segmentation Cues in Lexical Identification and in Lexical Acquisition: Same or Different?**

*Odile Bagou, Ulrich H. Frauenfelder, University of Geneva, Switzerland*

PAGE 2835  
FriSe3.P4-4

### **Phonological Representations in Poor Readers**

*Cecile Kuijpers, Louis ten Bosch, Radboud Universiteit Nijmegen, The Netherlands*

PAGE 2839  
FriSe3.P4-5

### **To What Extent Does Tagged-MRI Technique Allow to Infer Tongue Muscles' Activation Pattern? A Modelling Study**

*Stéphanie Buchaillard<sup>1</sup>, Pascal Perrier<sup>1</sup>, Yohan Payan<sup>2</sup>*  
<sup>1</sup>GIPSA, France; <sup>2</sup>TIMC-IMAG, France

PAGE 2843  
FriSe3.P4-6

### **Feature Adaptation of Hearing-Impaired Lip Shapes: The Vowel Case in the Cued Speech Context**

*Noureddine Aboutabit<sup>1</sup>, Denis Beautemps<sup>1</sup>, Olivier Mathieu<sup>1</sup>, Laurent Besacier<sup>2</sup>*  
<sup>1</sup>GIPSA, France; <sup>2</sup>LIG, France

PAGE 2847  
FriSe3.P4-7

### **Automatic Detection of the Context of Acoustic Landmark Deletion**

---

*FriSe3.P4 continued ...*

PAGE 2851  
FriSe3.P4-8

### **Aspects of Pharyngealized Phonemes in Arabic Using Articulography**

*Slim Ouni, LORIA, France*

PAGE 2852  
FriSe3.P4-9

### **The Effect of Spectral Tilt on Infants' Discrimination of Fricatives**

*Elizabeth Beach<sup>1</sup>, Christine Kitamura<sup>1</sup>, Harvey Dillon<sup>2</sup>, Teresa Ching<sup>2</sup>, Denis Burnham<sup>1</sup>*

<sup>1</sup>University of Western Sydney, Australia; <sup>2</sup>National Acoustic Laboratories, Australia

PAGE 2853  
FriSe3.P4-10

### **“Look at the Shark”: Evaluation of Student Produced Standardized Sentences of Infant- and Foreigner-Directed Speech**

*Monja Knoll, Lisa Scharrer, University of Portsmouth, UK*

PAGE 2857  
FriSe3.P4-11

### **Vocal Tract Inversion by Cepstral Analysis-by-Synthesis Using Chain Matrices**

*Sankaran Panchapagesan, Abeer Alwan, University of California at Los Angeles, USA*

PAGE 2861  
FriSe3.P4-12

### **DC-Constrained Linear Prediction for Glottal Inverse Filtering**

*Paavo Alku, Carlo Magi, Tom Bäckström, Helsinki University of Technology, Finland*

PAGE 2865  
FriSe3.P4-13

### **Voicing Influences the Saliency of Place of Articulation in Audio-Visual Speech Perception in Babble**

*Magnus Alm, Dawn Behne, NTNU, Norway*

PAGE 2869  
FriSe3.P4-14

### **Correspondence of Perception and Production Boundaries Between Single and Geminate Stops in Japanese**

*Shigeaki Amano<sup>1</sup>, Yukari Hirata<sup>2</sup>*

<sup>1</sup>NTT Corporation, Japan; <sup>2</sup>Colgate University, USA

PAGE 2873  
FriSe3.P4-15

### **Inhibitory Processes of Chinese Spoken Word Recognition**

*Michael C.W. Yip, Hong Kong Institute of Education, China*