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.

Printed by Curran Associates, Inc. (2009)

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

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
Web: www.proceedings.com

TueSe3.O2 : Speech Coding

Plaza 1, Time 13:30 – 15:30, Tuesday 23rd September 2008 — *Chair: Julien Epps*

TueSe3.O3 : Human Conversation and Communication

Plaza 2, Time 13:30 – 15:30, Tuesday 23rd September 2008 — *Chair: Bernd Möbius*

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*

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*

TueSe3.P1 continued …

TueSe3.P1 continued …

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*

TueSe3.P3 : Spoken Language Systems I

Mezzanine Level Area B3, Time 13:30 – 15:30, Tuesday 23rd September 2008 *Chair: Sebastian Möller*

TueSe3.P3 continued …

TueSe3.P4 : Emotion and Expression I

Mezzanine Level Area B4, Time 13:30 – 15:30, Tuesday 23rd September 2008 *Chair: Denis Burnham*

TueSe3.P4 continued …

Detection of Security Related Affect and Behaviour in Passenger Transport

Björn Schuller[1], *Matthias Wimmer*[2], *Dejan Arsic*[1], *Tobias Moosmayr*[3], *Gerhard Rigoll*[1]
[1] *Technische Universität München, Germany;* [2] *Waseda University, Japan;* [3] *BMW Group, Germany*

TueSe4.O1 : Automatic Speech Recognition: Acoustic Models I

Great Hall, Time 16:00 – 18:00, Tuesday 23rd September 2008 *Chair: Roger K. Moore*

TueSe4.O2: Accent and Language Identification

Plaza 1, Time 16:00–18:00, Tuesday 23rd September 2008 — *Chair: Yuko Kinoshita*

TueSe4.O3 : Emotion and Expression II

Plaza 2, Time 16:00 – 18:00, Tuesday 23rd September 2008 — *Chair: Ailbhe Ní Chasaide*

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*

TueSe4.P1 : Speaker Recognition and Diarisation

Mezzanine Level Area A1, Time 16:00 – 18:00, Tuesday 23rd September 2008 — *Chair: Robbie Vogt*

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*

TueSe4.P2 continued …

TueSe4.P2 continued …

TueSe4.P3 : Spoken Language Systems II

Mezzanine Level Area B3, Time 16:00 – 18:00, Tuesday 23rd September 2008 *Chair: Jerome Bellegarda*

TueSe4.P3 continued …

TueSe4.P3 continued …

TueSe4.P4: Perception, Production, Discourse and Dialog

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

TueSe4.P4 continued …

WedSe2.O1 : Single-Channel Speech Enhancement

Great Hall, Time 10:00 – 12:00, Wednesday 24th September 2008 — *Chair: Frank K. Soong*

WedSe2.O2: Speech Synthesis Methods I

Plaza 1, Time 10:00–12:00, Wednesday 24th September 2008 — *Chair: Rolf Carlson*

WedSe2.O3: Speaking Style and Emotion Recognition

Plaza 2, Time 10:00–12:00, Wednesday 24th September 2008 — *Chair: Javier Hernando*

WedSe2.O3 continued …

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*

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*

WedSe2.P1 continued …

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*

WedSe2.P2 continued …

WedSe2.P3 : Accent and Language Recognition

Mezzanine Level Area B3, Time 10:00 – 12:00, Wednesday 24th September 2008 — *Chair: Dirk Van Compernolle*

WedSe2.P3 continued …

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*

WedSe2.P4 continued …

WedSe3.O1 : Automatic Speech Recognition: Language Models I

Great Hall, Time 13:30 – 15:30, Wednesday 24th September 2008 *Chair: Keikichi Hirose*

WedSe3.O2 : Speaker Identification and Verification

Plaza 1, Time 13:30 – 15:30, Wednesday 24th September 2008 *Chair: Jean-François Bonastre*

WedSe3.O3: Prosodic Structure and Processing

Plaza 2, Time 13:30–15:30, Wednesday 24th September 2008 — *Chair: John Ingram*

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*

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*

WedSe3.P1 continued …

WedSe3.P2 : Robust Automatic Speech Recognition I

Mezzanine Level Area A2, Time 13:30 - 15:30, Wednesday 24th September 2008 — *Chair: Kuldip K. Paliwal*

WedSe3.P2 continued …

WedSe3.P2 continued …

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*

WedSe3.P3 continued …

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*

WedSe3.P4 continued …

WedSe4.O1 : Automatic Speech Recognition: Tone Languages

Great Hall, Time 16:00 – 18:00, Wednesday 24th September 2008 *Chair: Björn Granström*

WedSe4.O2 : Spoken Dialogue Systems

Plaza 1, Time 16:00 – 18:00, Wednesday 24th September 2008 — *Chair: Isabel Trancoso*

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

Plaza 2, Time 16:00–18:00, Wednesday 24th September 2008 — *Chair: Catherine T. Best*

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*

WedSe4.P1 : Automatic Speech Recognition: Adaptation I

Mezzanine Level Area A1, Time 16:00 – 18:00, Wednesday 24th September 2008 *Chair: Hervé Bourlard*

WedSe4.P1 continued …

WedSe4.P2 : Robust Automatic Speech Recognition II

Mezzanine Level Area A2, Time 16:00 – 18:00, Wednesday 24th September 2008 *Chair: Satoshi Nakamura*

WedSe4.P2 continued …

WedSe4.P2 continued …

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*

WedSe4.P3 continued …

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*

WedSe4.P4 continued …

WedSe4.P4 continued …

ThuSe2.O1 : Broadcast Transcription Systems

Great Hall, Time 10:00 – 12:00, Thursday 25th September 2008 — *Chair: Tanja Schultz*

ThuSe2.O2 : Voice Conversion and Modification

Plaza 1, Time 10:00 – 11:40, Thursday 25th September 2008 *Chair: Chris Davis*

ThuSe2.O3 : Phonetics: General

Plaza 2, Time 10:00 – 11:40, Thursday 25th September 2008 *Chair: Andrew Butcher*

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*

ThuSe2.P1 : Automatic Speech Recognition: Features I

Mezzanine Level Area A1, Time 10:00 – 12:00, Thursday 25th September 2008 — *Chair: Lin-shan Lee*

ThuSe2.P1 continued ...

ThuSe2.P2 : Automatic Speech Recognition: Language Models II

Mezzanine Level Area A2, Time 10:00 – 12:00, Thursday 25th September 2008 *Chair: Günther Ruske*

ThuSe2.P2 continued …

ThuSe2.P2 continued …

ThuSe2.P3 : Speech Resources and Technology Evaluation

Mezzanine Level Area B3, Time 10:00 – 12:00, Thursday 25th September 2008 *Chair: Christoph Draxler*

ThuSe2.P3 continued …

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*

ThuSe2.P4 continued …

Detection of Feeling Through Back-Channels in Spoken Dialogue

Tatsuya Kawahara[1], *Masayoshi Toyokura*[1], *Teruhisa Misu*[2], *Chiori Hori*[2]
[1]*Kyoto University, Japan;* [2]*NICT, Japan*

ThuSe3.O1 : Automatic Speech Recognition: Adaptation II

Great Hall, Time 13:30 – 15:30, Thursday 25th September 2008 *Chair: Sadaoki Furui*

ThuSe3.O2 : Applications in Education and Learning I

Plaza 1, Time 13:30 – 15:30, Thursday 25th September 2008 *Chair: Helmer Strik*

ThuSe3.O3 : Speech Pathologies

Plaza 2, Time 13:30 – 15:30, Thursday 25th September 2008 *Chair: Phil Green*

ThuSe3.O3 continued …

PAGE 1761
ThuSe3.O3-6
15:10 – 15:30

Application of Weighted Finite-State Transducers to Improve Recognition Accuracy for Dysarthric Speech

Omar Caballero Morales, Stephen Cox, University of East Anglia, UK

ThuSe3.O4: 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*

ThuSe3.O4 continued …

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*

ThuSe3.P2 : Speech Synthesis Methods II

Mezzanine Level Area A2, Time 13:30 – 15:30, Thursday 25th September 2008 *Chair: B. Yegnanarayana*

ThuSe3.P2 continued …

ThuSe3.P2 continued …

ThuSe3.P3 : Speaker Recognition: Adverse Conditions and Forensics

Mezzanine Level Area B3, Time 13:30 – 15:30, Thursday 25th September 2008 *Chair: Kiyohiro Shikano*

ThuSe3.P3 continued …

ThuSe3.P3 continued …

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*

ThuSe3.P4 continued …

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*

ThuSe4.O2 : Multimodal Signal Processing

Plaza 1, Time 16:00 – 18:00, Thursday 25th September 2008 — *Chair: Roland Göcke*

ThuSe4.O2 continued …

PAGE 2036
ThuSe4.O2-6
17:40 – 18:00

Feature Space Transforms for Czech Sign-Language Recognition

Jan Trmal, Marek Hrúz, 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: Jeesun Kim*

ThuSe4.O4: Evaluation and Standardisation of Spoken-Language Technology

Plaza 3&4, Time 16:00 – 18:00, Thursday 25th September 2008 *Chair: Bruce Millar*

ThuSe4.P1 : Automatic Speech Recognition: Search Methods

Mezzanine Level Area A1, Time 16:00 – 18:00, Thursday 25th September 2008 *Chair: Hugo Van hamme*

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*

ThuSe4.P2 continued …

Prosody Boundary Detection Through Context-Dependent Position Models

Yue-Ning Hu[1], *Min Chu*[2], *Chao Huang*[2], *Yan-Ning Zhang*[1]
[1]*Northwestern Polytechnical University, China;* [2]*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*

ThuSe4.P3 continued …

ThuSe4.P4 : Applications for the Aged and Handicapped

Mezzanine Level Area B4, Time 16:00 – 18:00, Thursday 25th September 2008 *Chair: Stephen Cox*

ThuSe4.P4 continued …

ThuSe4.P4 continued …

FriSe2.O1 : Automatic Speech Recognition: Features II

Great Hall, Time 10:00 – 12:00, Friday 26th September 2008 *Chair: Michael Wagner*

FriSe2.O2 : Speech Synthesis: Prosody and Emotion II

Plaza 1, Time 10:00 – 12:00, Friday 26th September 2008 *Chair: Masami Akamine*

FriSe2.O3 : Human Speech Production

Plaza 2, Time 10:00 – 12:00, Friday 26th September 2008 — *Chair: Janet Fletcher*

FriSe2.O4: 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*

FriSe2.O4 continued …

FriSe2.O5 : Spoken Language Translation Systems

Plaza 5, Time 10:00 – 12:00, Friday 26th September 2008 *Chair: Yuqing Gao*

FriSe2.P1 : Automatic Speech Recognition: Acoustic Models III

Mezzanine Level Area A1, Time 10:00 – 12:00, Friday 26th September 2008 — *Chair: Lori Lamel*

FriSe2.P1 continued …

FriSe2.P1 continued …

FriSe2.P2 : Spoken Language: Parsing and Summarisation

Mezzanine Level Area A2, Time 10:00 – 12:00, Friday 26th September 2008 — *Chair: Pietro Laface*

FriSe2.P2 continued …

FriSe2.P3 : Multimodal Interfaces

Mezzanine Level Area B3, Time 10:00 – 12:00, Friday 26th September 2008 *Chair: David House*

FriSe2.P3 continued …

FriSe2.P4 : Speech, Music, Audio Segmentation and Classification

Mezzanine Level Area B4, Time 10:00 – 12:00, Friday 26th September 2008 — *Chair: Haizhou Li*

FriSe2.P4 continued …

FriSe3.O1 : Automatic Speech Recognition: New Paradigms

Great Hall, Time 13:30 – 15:30, Friday 26th September 2008 *Chair: Alex Acero*

FriSe3.O2 : Speech and Acoustic Activity Detection

Plaza 1, Time 13:30 – 15:30, Friday 26th September 2008 *Chair: Jean-Pierre Martens*

FriSe3.O3 : Speech Analysis and Processing

Plaza 2, Time 13:30 – 15:30, Friday 26th September 2008 — *Chair: Catherine I. Watson*

FriSe3.O4: Special Session: Talking Heads and Pronunciation Training

Plaza 3&4, Time 13:30 – 15:30, Friday 26th September 2008 *Chair: Gérard Bailly*

FriSe3.O4 continued …

FriSe3.O5 : Multimodal Speech Processing

Plaza 5, Time 13:30 – 15:30, Friday 26th September 2008 *Chair: Takaaki Kuratate*

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*

FriSe3.P1 continued …

FriSe3.P1 continued …

FriSe3.P2 : Expression, Emotion and Personality Recognition

Mezzanine Level Area A2, Time 13:30 – 15:30, Friday 26th September 2008 *Chair: Kjell Elenius*

FriSe3.P3 : Applications in Education and Learning II

Mezzanine Level Area B3, Time 13:30 – 15:30, Friday 26th September 2008 *Chair: Gordon McIntyre*

FriSe3.P3 continued …

FriSe3.P4: Human Speech Production and Speech Perception

Mezzanine Level Area B4, Time 13:30–15:30, Friday 26th September 2008 — *Chair: Eva Hajičová*

FriSe3.P4 continued …