Welcome to the Seventh ISCA Tutorial and Research Workshop (ITRW) on Speech Synthesis!

This is the ISCA 7th Speech Synthesis Workshop. Twenty years have passed since the first one at Autrans, France in 1990, followed by Mohonk, USA in 1994, Jenolan Caves, Australia in 1998, Pitlochry, Scotland in 2001, Pittsburgh, USA in 2004, and Bonn, Germany in 2007.

Over the past twenty years, the quality of speech synthesis has been highly improved with the use of large speech data through corpus-based approach. The unit selection speech synthesis which was one of the tutorials in the first workshop has been widely employed and more flexible schemes have been enthusiastically pursued. In particular, statistical parametric speech synthesis based on hidden Markov models (HMMs) has been developed to cope with wider applications. So called paralinguistic information treatment has also been studied to synthesize not only reading speech but also communicative speech. For communicative speech synthesis, we still need huge speech data if we continue to use the current synthesis schemes. We invited Prof. Simon King of the University of Edinburgh as one of tutorial speakers. He will present a new step towards flexible speech synthesis without huge speech data collection.

The statistical parametric approach has provided a flexible framework for various applications such as multilingual and cross-lingual synthesis, on-line and mobile systems, conversational and expressive speech synthesis. Speaker characteristics control such as speaker adaptation and voice conversion has also been studied. Though synthesis model by itself is an important and tough problem, quite a few research efforts have been devoted to new modeling alternative to traditional source-filter models. We invited Prof. Hideki Kawahara of Wakayama University as the other tutorial speaker to provide us evolving synthesis scheme of STRAIGHT synthesis.

Furthermore, we have also planned "open source initiatives for speech synthesis" as a special event of this workshop, SSW7, on the last day in the evening to discuss open source activities in the community. We also would like to encourage all participants of SSW7 to attend Blizzard Challenge Workshop, which will be held just after SSW7 in the same place.

The workshop will be held as a satellite workshop of INTERSPEECH 2010 (Chiba, Japan, September 26-30, 2010). We selected ATR in Kyoto as the workshop site, where corpus-based approach was initiated. This workshop is co-sponsored by the National Institute of Information and Communications Technology (NICT), and the Effective Multilingual Interaction in Mobile
Environments (EMIME) project. We are grateful for their support.

We would like to express our sincere appreciation to members of the local organizing committee for their efforts for the preparation of the workshop. We are also thankful to the advisory committee members and program committee members for their advice and great effort in the review process.

Finally we welcome you to SSW7 again. We hope that you enjoy this workshop and benefit from the presentations and contributions.

Yoshinori Sagisaka
Keiichi Tokuda
Co-Chairs, SSW7
## Program at a glance

<table>
<thead>
<tr>
<th>Sept. 22</th>
<th>Sept. 23</th>
<th>Sept. 24</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00-10:00</td>
<td><strong>Registration</strong></td>
<td>9:00-10:40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lecture Session 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statistical Parametric Speech Synthesis</td>
</tr>
<tr>
<td>10:00-10:10</td>
<td><strong>Opening Remarks</strong></td>
<td></td>
</tr>
<tr>
<td>10:10-11:00</td>
<td><strong>Tutorial 1</strong></td>
<td>10:40-11:00</td>
</tr>
<tr>
<td></td>
<td>By Hidoki Kawahara</td>
<td><strong>Coffee Break</strong></td>
</tr>
<tr>
<td>10:40-11:00</td>
<td><strong>Coffee Break</strong></td>
<td></td>
</tr>
<tr>
<td>11:00-11:25</td>
<td><strong>Coffee Break</strong></td>
<td>11:00-11:50</td>
</tr>
<tr>
<td>11:25-12:40</td>
<td><strong>Lecture Session 1</strong></td>
<td><strong>Tutorial 2</strong></td>
</tr>
<tr>
<td></td>
<td>Concatenative Speech Synthesis</td>
<td>By Simon King</td>
</tr>
<tr>
<td>11:00-12:40</td>
<td><strong>Lecture Session 4</strong></td>
<td>11:50-12:40</td>
</tr>
<tr>
<td></td>
<td>Expressive Speech Synthesis</td>
<td>Lecture Session 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multi-Lingual Speech Synthesis</td>
</tr>
<tr>
<td>12:40-14:00</td>
<td><strong>Lunch</strong></td>
<td>12:40-14:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td>14:00-15:20</td>
<td><strong>Poster Session 1</strong></td>
<td>14:00-15:20</td>
</tr>
<tr>
<td></td>
<td>&amp; Coffee</td>
<td>Poster Session 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&amp; Coffee</td>
</tr>
<tr>
<td>14:00-15:20</td>
<td><strong>Poster Session 2</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&amp; Coffee</td>
<td></td>
</tr>
<tr>
<td>15:20-17:25</td>
<td><strong>Lecture Session 2</strong></td>
<td>15:20-17:00</td>
</tr>
<tr>
<td></td>
<td>Voice Conversion</td>
<td>Lecture Session 8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Selected Topics</td>
</tr>
<tr>
<td>15:20-17:00</td>
<td><strong>Lecture Session 5</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluation and Applications</td>
<td></td>
</tr>
<tr>
<td>17:30-18:00</td>
<td>move to Nara by bus</td>
<td>17:00-17:10</td>
</tr>
<tr>
<td>18:00-19:00</td>
<td>Free time in Nara for walking around and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shopping</td>
<td></td>
</tr>
<tr>
<td>17:25-17:40</td>
<td><strong>Message from SynSIG</strong></td>
<td></td>
</tr>
<tr>
<td>18:00-20:00</td>
<td><strong>Welcome Reception</strong></td>
<td>18:00-20:00</td>
</tr>
<tr>
<td></td>
<td>@ ATR Cafeteria (1\textsuperscript{st} floor)</td>
<td>Optional:</td>
</tr>
<tr>
<td>19:00-21:00</td>
<td><strong>Banquet</strong></td>
<td>Open Source Initiatives for Speech Synthesis</td>
</tr>
<tr>
<td></td>
<td>@ Nikko Nara Hotel (4\textsuperscript{th} floor)</td>
<td>@ Keihan Plaza Hotel (3\textsuperscript{rd} floor)</td>
</tr>
</tbody>
</table>
September 22, 2010

10:00-10:10 Opening Remarks

10:10-11:00 Tutorial 1
Chair: Yoshinori Sagisaka

T-1 Exploration of the other aspect of Vocoder revisited, -- A-Z STRAIGHT, TANDEM-STRAIGHT and morphing --
Hideki Kawahara

11:25-12:40 Lecture Session 1: Concatenative Speech Synthesis
Chair: Jerome Bellegarda

L-1.1 Crafting small data bases for unit selection TTS: effects on intelligibility
H. Timothy Bunnell (Center for Pediatric Auditory and Speech Sciences, Nemours Biomedical Research, USA)

L-1.2 Composite TTS voices
Alistair Conkie, and Ann K. Syrdal (AT&T Labs – Research Florham Park, NJ, USA)

L-1.3 Compression of line spectral frequency parameters using the asynchronous interpolation model
Alexander Kain and Todd Leen (Division of Biomedical Computer Science Oregon Health & Science University Portland, USA)

14:00-15:20 Poster Session 1 & Coffee
Chair: Heiga Zen

P-1.1 Implementation of VTLN for statistical speech synthesis
Lakshmi Saheer1,2, John Dines1, Philip N. Garner1, and Hui Liang1,2 ( 1 Idiap Research Institute, Martigny, 2 Switzerland Ecole Polytechnique Fédérale de Lausanne, Switzerland

P-1.2 Do prosodic cues influence uncertainty perception in articulatory speech synthesis?
Eva Lasarcyk1 and Charlotte Wollermann 2,3 (1Institute of Phonetics, Saarland University, Germany, 2 Institute of Communication Sciences, University of Bonn, Germany, 3 German Linguistics, University of Duisburg-Essen, Germany)

P-1.3 An unified and automatic approach of Mandarin HTS system
Yong Guan1, Jilei Tian1, Yi-Jian Wu2, Junichi Yamagishi3, and Jani Nurminen4 (1Nokia Research Center, Beijing, 2 Microsoft, China, 3 University of Edinburgh, UK 4 Nokia Devices R&D, Finland)

P-1.4 Synthesis of listener vocalisations with imposed intonation contours
Sathish Pamm1, Marc Schroeder2, Marcela Charfuluan1, Oytun Turk2, and Ingmar Steiner1 (1DFKI GmbH, Germany, 2 Sensory Inc., Portland, OR, USA)

P-1.5 An investigation of the impact of speech transcript errors on HMM voices
Jinfu Ni and Hisashi Kawai (Spoken Language Communication Group, MASTAR project, National Institute of Information and Communications Technology, Japan)

P-1.6 An HMM-based singing style modeling system for singing voice synthesizers
Keiji Saino, Makoto Tachibana, and Hideki Kenmochi (Corporate Research & Development Center, Yamaha Corporation, Japan)

P-1.7 Lombard effect mimicking
Dong-Yan Huang, Susanto Rahardja, and Ee Ping Ong (Signal Processing Department Institute for Infocomm Research)

P-1.8 Unsupervised prosody labeling for constructing Mandarin TTS
Chen-Yu Chiang, Sin-Horng Chen, and Yih-Ru Wang (Institute of Communication Engineering, National Chiao Tung University, Taiwan)

P-1.9 Analysis and synthesis of hypo and hyperarticulated speech
Benjamin Picart, Thomas Drugman, and Thierry Dutoit (TCTS Lab, Faculté Polytechnique (FPMs), University of Mons (UMons), Belgium)

P-1.10 Evaluating prosody in synthetic speech with online (eye-tracking) and offline (rating) methods
Rajakrishnan Rajkumar, Michael White, Shari R. Speer, and Kiwako Ito (Department of Linguistics, The Ohio State University, USA)
Lecture Session 2: Voice Conversion

Chair: Frank Soong

L-2.1 GMM-PCA based speaker-timbre conversion on full-quality speech
Fernando Villavicencio and Esteban Maestre (Music Technology Group, Universitat Pompeu Fabra, Spain)

L-2.2 Voice conversion using precise speech alignment based on spectral property and eigen-codeword distribution
Yi-Chin Huang, Chung-Hsien Wu, Chung-Han Lee, and Yu-Ting Chao (Department of Computer Science and Information Engineering, National Cheng Kung University, Taiwan)

L-2.3 On transforming spectral peaks in voice conversion
Elizabeth Godoy1, Olivier Rosec1, and Thierry Chonavel2 (1 Orange Labs R&D TECH/ASAP/VOICE, France, 2 Télécom Bretagne, Signal & Communications Department, France)

L-2.4 Linear transformation approaches to many-to-one voice conversion
Chie Hayashida, Tomoki Toda, Yamato Ohitani, Hiroshi Saruwatari, and Kiyohiro Shikano (Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), Japan)

L-2.5 HMM-based robust voice conversion using adaptive F0 quantization
Takashi Nose and Takao Kobayashi (Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Japan)

Message from SynSIG

September 23, 2010

Lecture Session 3: Statistical Parametric Speech Synthesis

Chair: Simon King

L-3.1 Statistical parametric speech synthesis with joint estimation of acoustic and excitation model parameters
Ranniery Maia, Heiga Zen, and Mark Gales (Toshiba Research Europe Ltd., Cambridge Research Laboratory, UK)

L-3.2 From discontinuous to continuous F0 modelling in HMM-based speech synthesis
Kai Yu, Blaise Thomson, and Steve Young (Cambridge University Engineering Department, UK)

L-3.3 Spectral modeling with contextual additive structure for HMM-based speech synthesis
Shinji Takaki, Yoshihiko Nankaku, and Keiichi Tokuda (Department of Computer Science and Engineering, Nagoya Institute of Technology, Japan)

L-3.4 Bayesian speech synthesis framework integrating training and synthesis processes
Kei Hashimoto, Yoshihiko Nankaku, and Keiichi Tokuda (Department of Scientific and Engineering Simulation, Nagoya Institute of Technology, Japan)

Lecture Session 4: Expressive Speech Synthesis

Chair: Nick Campbell

L-4.1 Symbolic vs. acoustics-based style control for expressive unit selection
Ingmar Steiner1,2, Marc Schroeder1, Marcela Charfuelan1, and Annette Klepp1,2 (1 DFKI GmbH, Germany, 2 Department of Computational Linguistics & Phonetics, Saarland University, Germany)

L-4.2 Application of expressive TTS synthesis in an advanced ECA system
Jan Romportl1, Enrico Zovato2, Raul Santos3, Pavel Iríng3, Jose Relano Gil1, and Morena Danieli2 (1 Department of Cybernetics, University of West Bohemia, Czech Republic, 2 Loquendo, S.p.A., Italy, 3 Telefónica I+D, Spain)

L-4.3 A hidden Markov model-based approach for emotional speech synthesis
Chih-Yung Yang and Chia-Ping Chen (Department of Computer Science and Engineering National Sun Yat-Sen University Kaohsiung, Taiwan)

L-4.4 Two vocoder techniques for neutral to emotional timbre conversion
Fabio Tesser1, Enrico Zovato2, Mauro Nicolao3, and Piero Cosi1 (1 Institute of Cognitive Sciences and Technologies, Italian National Research Council, Padova, Italy 2 Loquendo S.p.A., Italy)
Chair: Alistair Conkie

P-2.1  
**Refined statistical model tuning for speech synthesis**
Xu Shao, Vincent Pollet, and Andrew Breen (TTS R&D, Nuance Communications)

P-2.2  
**High quality TTS voices within one day**
Didier Cadic1 and Christophe d’Alessandro2 (1 Orange Labs, France, 2 CNRS-LIMSI, France)

P-2.3  
**Nativization of English words in Spanish using analogy**
Tatyana Polyakova and Antonio Bonafonte (Universitat Politècnica de Catalunya, Barcelona, Spain)

P-2.4  
**Automatic prosodic labeling of accent information for Japanese spoken sentences**
Asami Yamamoto, Kazuhiro Suzuki, Kook Cho, and Yoichi Yamashita (College of Information Science and Engineering, Ritsumeikan University, Japan)

P-2.5  
**An automatic pitch model with distance function**
Mohamed Abou-Zleikha, Peter Cahill, and Julie Carson-Berndsen (CNGL, School of Computer Science and Informatics, University College Dublin, Ireland)

P-2.6  
**Considering readability in Text-to-Speech recording script design**
Minghui Dong, Ling Cen, Paul Chan, and Haizhou Li (Human Language Technology Department, Institute for Infocomm Research, A*STAR, Singapore)

P-2.7  
**Letter-based speech synthesis**
Oliver Watts, Junichi Yamagishi, and Simon King (Centre for Speech Technology Research, University of Edinburgh, UK)

P-2.8  
**Joint prosodic and segmental unit selection for expressive speech synthesis**
Christophe Veaux, Pierre Lanchantin, and Xavier Rodet (IRCAM – CNRS STMS, Analysis-Synthesis Team, 1, France)

P-2.9  
**Speech synthesis in the mobile user interface**
Pieter E. Scholz1,2, Justus C. Roux1, and Jacques P. du Toit1 (1 Department of Electrical and Electronic Engineering, Stellenbosch University, South Africa, 2 CatchWord Language & Speech Technologies (Pty) Ltd, Stellenbosch, South Africa, 3 School of Languages, North-West University, Potchefstroom, South Africa)

Chair: Gerard Bailly

L-5.1  
**Evaluating speech synthesis intelligibility using Amazon Mechanical Turk**
Maria K. Wolters, Karl B. Isaac, and Steve Renals (Centre for Speech Technology Research, University of Edinburgh, United Kingdom)

L-5.2  
**Further exploration of the possibilities and pitfalls of multidimensional scaling as a tool for the evaluation of the quality of synthesized speech**
Anna C. Janska1 and Robert A.J. Clark2 (1 IMPRS NeuroCom, University of Leipzig, Germany, 2 CSTR, The University of Edinburgh, U.K.)

L-5.3  
**Handling large audio files in audio books for building synthetic voices**
Kishore Prahalad1,2 and Alan W Black1 (1 International Institute of Information Technology, Hyderabad, India, 2 Language Technologies Institute, Carnegie Mellon University, USA)

L-5.4  
**Improving speech synthesis for noisy environments**
Gopala Krishna Anumanchipalli, Prasanna Kumar Muthukumar, Udhyakumar Nallasamy, Alok Parlikar, Alan W Black, and Brian Langner (Language Technologies Institute Carnegie Mellon University, USA)

September 24, 2010

Chair: Paul Taylor

L-6.1  
**Learning speaker-specific phrase breaks for Text-to-Speech systems**
Kishore Prahalad1,2, E. Veeva Raghavendra1, and Alan W Black1 (1 International Institute of Information Technology, India, 2 Language Technologies Institute, Carnegie Mellon University, USA)

L-6.2  
**Substitution of state distributions to reproduce natural prosody on HMM-based speech synthesizers**
Nobuyuki Nishizawa and Tsuneo Kato (KDDI R&D Laboratories Inc., Japan)

L-6.3  
**Utilising spontaneous conversational speech in HMM-based speech synthesis**
Sebastian Andersson, Junichi Yamagishi, and Robert Clark (The Centre for Speech Technology Research, University of
Lecture Session 7: Multi-Lingual Speech Synthesis

Chair: Alan Black

L-7.1 HMM-based polyglot speech synthesis by speaker and language adaptive training
Heiga Zen, Norbert Braunschweiler, Sabine Buchholz, Kate Knill, Sacha Krstulovic, and Javier Latorre (Toshiba Research Europe Ltd., Cambridge Research Laboratory, UK)

L-7.2 Speaker adaptation and the evaluation of speaker similarity in the EMIME speech-to-speech translation project
Mirjam Wester1, John Dines2, Matthew Gibson3, Hui Liang2, Yi-Jian Wu3, Lakshmi Saheer3, Simon King3, Keichiro Oura5, Philip N. Garner2, William Byrne3, Yong Guan6, Teemu Hirsimäki3, Reima Karhila4, Mikko Kurimo4, Matt Shannon6, Sayaka Shiota2, Jieli Tian1, Keichi Tokuda1 and Junichi Yamagishi1 (1 University of Edinburgh, UK, 2 Idiap Research Institute, Switzerland, 3 Aalto University, Finland, 4 University of Cambridge, UK, 5 Nagoya Institute of Technology, Japan, 6 Nokia Research Center Beijing, China)

Chair: Kishore Prabhallad

P-3.1 Comparison of formant enhancement methods for HMM-based speech synthesis
Tuomo Raitio1, Antti Suni2, Hannu Pulakka1, Martti Vainio2, and Paavo Älkü1 (1 Department of Signal Processing and Acoustics, Aalto University, Finland, 2 Department of Speech Sciences, University of Helsinki, Finland)

P-3.2 EM-HTS: real-time HMM-based Malay emotional speech synthesis
Mumtaz B. Mustafa, Raja N. Ainon, and Roziati Zainuddin (Faculty of Computer Science and Information Technology, University of Malaya, Malaysia)

P-3.3 High level emotional speech morphing using STRAIGHT
Dong-Yan Huang, Susanto Rahardja, and Ee Ping Ong (Institute for Infocomm Research, Singapore)

P-3.4 Adding speaking style to a TTS system
Jean-Philippe Goldman1,2, Sophie Roekhaut1,4, and Anne Catherine Simon2 (1 Department of Linguistics, University of Geneva, Switzerland, 2 Institut Langage & Communication/Valibèl – Discours & Variation, Université catholique de Louvain, Belgium, 3 TCTS Lab, University of Mons - UMONS, Belgium, 4 CENTAL, Université catholique de Louvain, Belgium)

P-3.5 Synthesizing fast speech by implementing multi-phone units in unit selection speech synthesis
Donata Moers1,2, Igor Jauk1, Bernd Mönich2,3, and Petra Wagner2 (1 Division of Language and Speech Communication, University of Bonn, Germany, 2 Fakultat für Linguistik und Literaturwissenschaft, University of Bielefeld, Germany, 3 IMS, University of Stuttgart, Germany)

P-3.6 Improved generation of prosodic features in HMM-based Mandarin speech synthesis
Miaomiao Wang1, Miaomiao Wen1, Daisuke Saito1, and Keikichi Hirose2, Nobuaki Minematsu2 (1 Graduate School of Engineering, The University of Tokyo, Japan, 2 Graduate School of Information Science and Technology, The University of Tokyo, Japan)

P-3.7 An HMM-based speech synthesiser using glottal post-filtering
João P. Cabral1,2, Steve Renals2, Korin Richmond2, and Junichi Yamagishi2 (1 School of Computer Science and Informatics, University College Dublin, Ireland 2 The Centre for Speech Technology Research, University of Edinburgh, UK)

P-3.8 A study of lexical stress patterns in unit selection synthesis
Yeon-Jun Kim and Mark C. Beutnagel (AT&T Labs - Research, USA)

P-3.9 Automatic prominence annotation of a German speech synthesis corpus: towards prominence-based prosody generation for unit selection synthesis
Andreas Windmann1, Petra Wagner2, Fabio Tamburini2, Denis Arnold1, and Catharine Oertel1 (1 Faculty of Linguistics and Literature, University of Bielefeld, Germany, 2 Department of Linguistics and Oriental Studies, University of Bologna, Italy, 3 Language and Speech Communication, University of Bonn, Germany)
### Lecture Session 8: Selected Topics

**Chair:** Bernd Moebius

**L-8.1** Toward naturally expressive speech synthesis: data-driven emotion detection using latent affective analysis  
Jerome R. Bellegarda (Speech & Language Technologies, Apple Inc., USA)

**L-8.2** KLATTSTAT: knowledge-based parametric speech synthesis  
Gopala Krishna Anumanchipalli¹, Ying-Chang Cheng², Joseph Fernandez², Xiaohan Huang², Qi Mao², and Alan W Black¹  
¹Language Technologies Institute, ²Electrical and Computer Engineering Carnegie Mellon University, USA

**L-8.3** Recent development of the HMM-based singing voice synthesis system – Sinsy  
Keiichiro Oura, Ayami Mase, Tomohiko Yamada, Satoru Muto, Yoshihiko Nankaku, and Keiichi Tokuda (Department of Computer Science, Nagoya Institute of Technology, Japan)

**L-8.4** Photo-real lips synthesis with trajectory-guided sample selection  
Lijuan Wang¹, Xiaojun Qian², Wei Han³, and Frank K. Soong¹  
¹Microsoft Research Asia, China, ²Department of Systems Engineering, Chinese University of Hong Kong, China, ³Department of Computer Science, Shanghai Jiao Tong University, China

---

**17:00-17:10** Closing Remarks
Committees

Co-Chairs
* Yoshinori Sagisaka, Waseda Univ., Japan
* Keiichi Tokuda, Nagoya Inst. of Tech., Japan

Local Organizing Committee
* Committee Chair
  Hisashi Kawai, NICT, Japan
* Committee Members
  Takehiko Kagoshima, Toshiba, Japan
  Shinsuke Sakai, NICT, Japan
  Yoshinori Shiga, NICT, Japan
  Tomoki Toda, NAIST, Japan
  Yoichi Yamashita, Ritsumeikan Univ., Japan

Advisory Committee
* Gerard Bailly, CNRS/INPG, France
* Alan Black, CMU, USA
* Nick Campbell, Univ. of Dublin, Ireland
* Rolf Carlson, KTH, Sweden
* Thierry Dutoit, Univ. of Mons, Belgium
* Wolfgang Hess, Univ. of Bonn, Germany
* Keikichi Hirose, Univ. of Tokyo, Japan
* Julia Hirschberg, Columbia Univ., USA
* Mari Ostendorf, Univ. of Washington, USA
* Jan van Santen, OHSU, USA
* Juergen Schroeter, AT&T Labs, USA
* Minoru Tsuzaki, Kyoto City Univ. of Arts, Japan

Program Committee
* Committee Chair
  Yoichi Yamashita, Ritsumeikan Univ., Japan
* Committee Members
  Masami Akamine, Toshiba, Japan
  Matthew Aylett, Cereproc, UK
  Gerard Bailly, CNRS/INPG, France
  Jerome Bellegarda, Apple, USA
  Alan Black, CMU, USA
  Andrew Breen, Nuance Communications, USA
  Nick Campbell, Univ. of Dublin, Ireland
  Alistair Conkie, AT&T Labs, USA
  Thierry Dutoit, Univ. of Mons, Belgium
  Raul Fernandez, IBM, USA
  Simon King, Univ. of Edinburgh, UK
  Takao Kobayashi, Tokyo Inst. of Tech., Japan
  Bernd Moehl, Univ. of Bonn, Germany
  Jinfu Ni, NICT, Japan
  Nobuyuki Nishizawa, KDDI, Japan
  Mari Ostendorf, Univ. of Washington, USA
  Kishore Prahallad, IIT, India
  Marc Schroeder, DFKI, Germany
  Frank Soong, Microsoft Research Asia, China
  Yannis Stylianou, Univ. of Crete, Greece
  David Talkin, Google, USA
  Jianhua Tao, Chinese Academy of Sciences, China
  Paul Taylor, Phonetic Arts, UK
  Minoru Tsuzaki, Kyoto City Univ. of Arts, Japan
Sponsors