Thursday 26 September - Volume 3

Keynote 5
Time and Place: 08.40 - 09.10 hrs Liguria Hall

Up from trigrams! The struggle for improved language model
— F. Jelinek, IBM T.J. Watson Research Center, USA

Keynote 6
Time and Place: 08.40 - 09.10 hrs Riviera Hall

Synthesis: Modelling variability and constraints
— R. Carlson, KTH, Sweden

Session 36: Dialogue and translation
Time and Place: 09.20 - 11.00 hrs Liguria Hall
Chairperson: H. Fujisaki, University of Tokyo, Japan

36.1 The role of dialogue in speech recognition, the case of the yellow pages system
— M. Guyomard, J. Siroux, A. Cozannet, ENSSAT-IRISA, IUT-IRISA, CNET, France

36.2 Interpretation of context-dependent utterances in man-machine dialogue
— E. Gerbino, P. Baggia, CSELT, Italy

36.3 The description of minor clauses in information-seeking telephone dialogues
— S. Eggins, J. Vonwiller, C.M.I. Matthiessen, P. Sefton, University of Sidney, Australia

36.4 Toward a spoken language translator for restricted-domain context-free languages
— D.B. Roe, F. Pereira, R. Sproul, M.D. Riley, P.J. Moreno, A. Macarron, AT&T Bell Labs., USA, Telefonica I + D, Spain

36.5 Bidirectional machine translation in Indian languages

Session 37: Speech analysis and signal representation
Time and Place: 09.20 - 11.00 hrs Riviera Hall
Chairperson: G. Chollet, CNRS-URA, France

37.1 Exact monitoring of the numerical error in various speech algorithms
— C. Papaodysseus, E. Koukoutsis, C. Triantafyllou, C. Vasilatos, National Technical University of Athens, Greece

37.2 Automatic computation and comparison of dynamically varying voice source parameters
— J. Koreman, B. Cranen, L. Boves, University of Nijmegen, The Netherlands

37.3 Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering
— P. Alku, Helsinki University of Technology, Finland

37.4 Generalized functional approximation for source-filter system modeling
— T. Galas, X. Rodet, Laforia-UA CNRS, France
Session 42: Algorithms for automatic speech recognition
Time and Place: 11.20 - 13.00 hrs
Chairperson: E. Frangoulis, Logica Cambridge, UK

42.1 A fast algorithm for deleted interpolation
— L.R. Bahl, P. Brown, P.V. de Souza, R.L. Mercer, D. Nahamoo, IBM Thomas J. Watson Research Center, USA

42.2 Recent work in continuous speech recognition using the connectionist Viterbi training procedure
— M.A. Franzini, A.H. Waibel, K.F. Lee, Telefonica I+D, Spain, Carnegie Mellon University, Apple Computer Corp., USA

42.3 A search organization for large-vocabulary recognition based on N-best decoding
— V. Steinbiss, Philips Research Lab. Aachen, Germany

42.4 VINICS: A continuous speech recognizer based on a new robust formulation
— Y. Gong, J.P. Haton, CRIN/CNRS/INRIA, France

42.5 A matrix representation of HMM-based speech recognition algorithms
— S. Sagayama, ATR Interpreting Telephony Res. Labs., Japan

Session 43: Text-to-speech synthesis systems
Time and Place: 11.20 - 13.00 hrs
Chairperson: C. Sorin, CNET, France

43.1 Speech Maker: text-to-speech conversion based on a multi-level, synchronized data structure
— E. te Lindert, H.C. van Leeuwen, Institute for Phonetics Sciences, TNO Institute for Perception Research, The Netherlands

43.2 A new text-to-speech synthesis system
— E. Lewis, M.A.A. Tatham, University of Bristol, University of Essex, UK

43.3 DIXI - Portuguese text-to-speech system
— L.C. Oliveira, M. Céu Viana, I.M. Trancoso, INESC/IST, CLUL, Portugal

43.4 Higher-level linguistic information in a text-to-speech system for Danish
— P. Molback Hansen, N. Reinholt Petersen, J. Rischel, C. Henriksen, Telecommunications Research Laboratory, Denmark

43.5 Adaptation of the Multivox text-to-speech system to Italian
— G. Olaszy, Hungarian Academy of Sciences, Hungary

Session 44: Phonetic modelling II
Time and Place: 11.20 - 13.00 hrs
Chairperson: V.W. Zue, MIT, USA

44.1 Correlation analysis of vowels and their application to speech recognition
— P. Niyogi, V.W. Zue, Massachusetts Institute of Technology, USA

44.2 Use of phonetic knowledge when designing and training stochastic models for speech recognition
— J.N. Holmes, Consultant, UK

44.3 Modelling phones by microsegments in a phonetically oriented recognition system
— B. Kaspar, K. Schuhmacher, Forschungsinstitut der DBP Telekom, Germany

44.4 An extended LVQ2 algorithm and its application to phoneme classification
— H.K. Kim, H.S. Lee, Korea Advanced Inst. of Science & Technology, Korea
44.5 A hierarchical broad phonetic classification scheme
— P.J. Dix, G.J. Vernooij, G. Bloothooft, University of Utrecht, The Netherlands

Session 45: Generation of prosody
Time and Place: 11.20 - 13.00 hrs
Chairperson: R. Carlson, KTH, Sweden

45.1 Using text analysis to predict intonational boundaries
— J. Hirschberg, AT&T Bell Labs., USA

45.2 Why do speakers accent "given" information?
— M. Horne, University of Lund, Sweden

45.3 Automatic prosody assignment for interactive synthesized dialogue systems
— J.P. Vonwiller, R.W. King, R.W.T. Lloyd, University of Sidney, Australia

45.4 Generating intonation in a voice dialogue system
— N. Youd, J. House, Logica Cambridge, University College London, UK

45.5 Computing linguistic knowledge for text-to-speech systems with PROSO
— R. Delmonte, R. Dolci, University of Venice, Italy

Session 46: Speech processing and analysis
Poster Session 14
Time and Place: 11.20 - 13.00 hrs
Chairperson: J.P. Lefevre, OROS, France

46.1 Acoustic echo cancellation using prediction residual signals
— C. Acker, P. Vary, H. Ostendarp, Technical University of Aachen, Germany

46.2 An evaluation of adaptive noise cancelling for speech recognition
— H.S. Dabis, A. Wrench, Paisley College, University of Edinburgh, UK

46.3 An efficient algorithm for real-time voiced/unvoiced decision
— E. Mumolo, A. Riccio, G. Abbattista, University of Trieste, Alcatel Face, Italy

46.4 Models of pitch perception
— T. Aarset, B. Gold, MIT Lincoln Laboratory, USA

46.5 A modified frequency warping function for PLP analysis
— Y. Gu, J.S. Mason, King’s College, University College of Swansea, UK

46.6 A new perspective on LPC excitation using singular value decomposition
— P. Corney, J.S. Mason, University of Wales, UK

46.7 Intra-speaker transplantation of speech characteristics: an application of waveform vocoding techniques and DTW
— W. Verhelst, M. Borger, Vrije Universiteit Brussel, Belgium, Eindhoven University of Technology, The Netherlands

46.8 Non linear dynamical aspects of speech waveforms
— T. Bullen, L. Condie, University of South Australia, Australia

46.9 Real-time signal processing for research, education and entertainment
— D. Talkin, AT&T Bell Laboratories, USA

46.10 Decomposition of the LPC excitation using wavelet functions
— S.H. Leung, O.Y. Wong, K.L. Laj, City Polytechnic of Hong Kong

46.11 An adaptive cochlear model for speech recognition
— E. Ambikairajah, L. Kilmartin, Regional Technical College, Ireland
46.12 Speech segmentation and classification using higher order moments
— G. Jacovitti, A. Falaschi, P. Pierucci, University of Rome, IBM Semea, Italy

Session 47: Automatic speech recognition: hardware and noise reduction
Poster Session 15
Time and Place: 11.20 - 13.00 hrs
Chairperson: A. Ciaramella, CSRLT, Italy

47.1 A PC - housed speaker independent large vocabulary continuous telephonic speech recognizer
— A. Ciaramella, D. Clementino, R. Pacifici, CSELT, Italy

47.2 Speaker independent continuous HMM-based recognition of isolated words on a real-time multi-DSP system
— A. Aktas, K. Zünkler, Siemens AG, Germany

47.3 A real time speech decoder using instantaneous frequency and energy
— A. Tsopanoglou, E.D. Kyriakis-Bitaros, J. Mourjopoulos, G. Kokkinakis, University of Patras, Greece

47.4 Fast hardware for efficient parallel processing of speech signals
— M. Schultheis, A. Lacroix, Institut für Angewandte Physik, Germany

47.5 The one chip speech recognition system

47.6 Influence of the telephone line on automatic speech recognition
— L. Villarrubia, M.J. Poza, C. Crespo, Telefónica I+D, Spain

47.7 Compensation for the effect of the communication channel in auditory-like analysis of speech (RASTA-PLP)
— H. Hermansky, A. Bayya, N. Morgan, P. Kohn, US West Advanced Technologies, International Computer Science Institute, USA

47.8 A study of endpoint detection algorithms in adverse conditions: incidence on a DTW and HMM recognizer
— J.C. Junqua, B. Reaves, B. Mak, Speech Technology Laboratory, USA

47.9 High-performance speech recognition in noise by continuously updated reference templates
— S. Dvorak, T. Hörmann, SEL Alcatel, Germany

47.10 Speech enhancement in the case of speech recognizers
— K. Vicsi, G. Békesy Acoustic Res. Lab. of Hungarian Academy of Sciences, Hungary

47.11 A robust feature extraction method for automatic speech recognition in noisy environments
— J. Gomez-Mena, J. Santos-Suarez, R. Garcia-Gomez, ETSI Telecomunicacion, Spain

Break for lunch: 13.00 - 14.00 hrs
Session 48: Sub-word units for automatic speech recognition
Time and Place: 14.00 - 15.40 hrs
Chairperson: R. Pieraccini, AT&T Bell Labs., USA

48.1 Selection of speech units for a speaker-independent CRS task
— L. Fissore, E. Giachin, G. Micca, P. Laface, CSELT, Polytechnic of Turin, Italy

48.2 Word juncture modeling using inter-word context-dependent phone-like units
— E. Giachin, C.H. Lee, L.R. Rabiner, A.E. Rosenberg, R. Pieraccini, CSELT, Italy, AT&T Bell Labs., USA

48.3 Phoneme-context-dependent LR parsing algorithms for HMM-based continuous speech recognition
— A. Nagai, S. Sagayama, K. Kita, ATR Interpreting Telephony Res. Labs., Japan

48.4 Optimizing lexical fast search in a large vocabulary isolated word speech recognition system
— H. Drexler, R. Roddeman, L. Boves, H. Strik, University of Nijmegen, The Netherlands

Session 49: Neural nets II
Time and Place: 14.00 - 15.40 hrs
Chairperson: F. Fallside, University of Cambridge, UK

49.1 Integrated phoneme-function word architecture of hidden control neural networks for continuous speech recognition
— B. Petek, A.H. Waibel, J.M. Tebelskis, Carnegie Mellon University, USA

49.2 Multiple dynamic features to enhance neural net based speaker verification
— X. Zhang, J.S. Mason, E.C. Andrews, University College of Swansea, UK

49.3 Time-delay neural networks embedding time alignment: a performance analysis
— P. Haffner, A. Waibel, CNET, France, Carnegie Mellon University, USA

49.4 Phoneme recognition using recurrent neural networks
— Y. Fukuda, H. Matsumoto, Kobe University, Japan

49.5 An integration of knowledge and neural networks toward a phoneme typewriter without a language model
— Y. Komori, K. Hatazaki, ATR Interpreting Telephony Res. Labs., NEC Corp., Japan

Session 50: Auditory modelling
Time and Place: 14.00 - 15.40 hrs
Chairperson: J.M. Dolmazon, I.C.P., France

50.1 Signal processing using an auditory filter bank with side-lobes and phase-jumps
— T. Fjällbrant, F. Mekuria, University of Linköping, Sweden

50.2 Notes on auditive coding of sophisticated signals
— J.S.C. van Dijk, University of Amsterdam, The Netherlands

50.3 An auditorily based spectral transformation of speech signals
— M. Beham, Technical University of Munich, Germany

50.4 On and off units detect information bottle-necks for speech recognition
— A.C. Morris, P. Escudier, J.L. Schwartz, Université Stendhal, France

XXXIII
### Session 51: Perception II

**Time and Place:** 14.00 - 15.40 hrs  
**Room A**  
*Chairperson: M. Wajskop, Université Libre de Bruxelles, Belgium*

<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Speaker(s)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>51.1</td>
<td>Detection times for vowels versus consonants</td>
<td>B. van Ooyen, A. Cutler, D. Norris, MRC Applied Psychology Unit, UK</td>
<td>Room A</td>
</tr>
<tr>
<td>51.2</td>
<td>The influence of sentence accent, word stress, and word class on the quality of vowels</td>
<td>D.R. van Bergem, University of Amsterdam, The Netherlands</td>
<td>Room A</td>
</tr>
<tr>
<td>51.3</td>
<td>A peak-and-level model for focus words in read and spontaneous natural speech and in synthetic speech</td>
<td>F.J. Koopmans-van Beinum, University of Amsterdam, The Netherlands</td>
<td>Room A</td>
</tr>
<tr>
<td>51.4</td>
<td>Connected speech processes in second language learning</td>
<td>J. Ingram, J. Pittam, University of Queensland, Australia</td>
<td>Room A</td>
</tr>
</tbody>
</table>

### Session 52: Speech interfaces: dialogue and human factors

**Poster Session 16**  
**Time and Place:** 14.00 - 15.40 hrs  
**Sector 1**  
*Chairperson: J. Laver, University of Edinburg, UK*

<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Speaker(s)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>52.1</td>
<td>Speech understanding and dialogue over the telephone: an overview of progress in the SUNDIAL project</td>
<td>J. Peckham, Logica Cambridge, UK</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.2</td>
<td>A system for natural spoken language queries: design, implementation and assessment</td>
<td>J.P. Tubach, P. Doignon, Télécom Paris, France</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.3</td>
<td>Operational validation of syntactic-semantic models in a spoken man-machine dialogue system</td>
<td>G. Deville, P. Mousel, Facultés Universitaires de Namur, Belgium, CRP-CU, Luxembourg</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.4</td>
<td>References in a multimodal dialogue: towards a unified processing</td>
<td>B. Gaiffe, L. Romary, J.M. Pierrel, CRIN/CNRS &amp; INRIA Lorraine, France</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.5</td>
<td>The user-Unix dialogue: a novel integrated approach to enhancing the operating system interface</td>
<td>P. Lefebvre, G. Duncan, F. Poirier, Siemens-Nixdorf Information Systems, Télécom Paris, France</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.6</td>
<td>Adoption of verbal and visual dialogue behaviour in document handling systems</td>
<td>B. Arndt, University of Stuttgart, Germany</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.7</td>
<td>The contribution of vision to speech perception</td>
<td>P.M.T. Smeele, A.C. Sittig, Delft University of Technology, The Netherlands</td>
<td>Sector 1</td>
</tr>
<tr>
<td>52.8</td>
<td>Processing disfluent speech: how and when are disfluencies found?</td>
<td>R.J. Lickley, R.C. Shillcock, E.G. Bard, University of Edinburgh, UK</td>
<td>Sector 1</td>
</tr>
</tbody>
</table>
52.9 Building a user interface for speech recognition - based telephone application system
— A. Choinière, R. Descout, J.M. Robert, Canadian Workplace Automation Research Centre - Communications, Ecole Polytechnique de Montréal, Canada

52.10 System design and human factors in auditory interfaces
— A.C. Murray, D.M. Jones, C.R. Frankish, University of Wales, Bristol University, UK

Coffee break: 15.40 - 16.00 hrs

PANEL SESSION 1
Industrial and Commercial Issues in Speech Technology
Time and Place: 16.00 - 17.30 hrs
Chairpersons:
P. Van Hove (CEC, Belgium)
J. Peckham (Logica Cambridge, UK)
Panelists:
J. Baker (Dragon Systems, USA)
H. Bourlard (Lernout & Hauspie Speechproducts, Belgium)
S. Furui (NTT Human Interface Labs., Japan)
H. Høge (Siemens, Germany)
F. Jelinek (IBM, USA)
J.P. Lefevre (OROS, France)
M.J. Hunt (Marconi, UK)
A. Macarron (Telefonica I + D, Spain)
J.C. Maurel (Infovox, Sweden)
A. Riccio (Face Alcatel, Italy)

PANEL SESSION 2
Interactions Between Speech and Language Processing
Time and Place: 16.00 - 17.30 hrs
Chairpersons:
J. Soler (CEC, Luxembourg)
H. Niemann (University of Erlangen, Germany)
Panelists:
L. Boves (University of Nijmegen, The Netherlands)
R. De Mori (McGill University, Canada)
F. Fallside (University of Cambridge, UK)
H. Fujisaki (University of Tokyo, Japan)
J.P. Haton (CRIN-INRIA, France)
C. Rullent (CSELT, Italy)
W. Von Hahn (University of Hamburg, Germany)
S. Young (CMU, USA)
V. Zue (MIT, USA)