### Technical Programme

<table>
<thead>
<tr>
<th>Volume 1</th>
<th>Pages K1 - K12, 1 - 744</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume 2</td>
<td>Pages 745 - 1518</td>
</tr>
<tr>
<td>Volume 3</td>
<td>Pages 1519 - 2270</td>
</tr>
<tr>
<td>Volume 4</td>
<td>Pages 2271 - 2844</td>
</tr>
</tbody>
</table>

Affiliations may have been abbreviated

Notation applied in Session allocation: \( X_{n_1n_2} \)

- **X = Day**
  - A: Monday
  - B: Tuesday
  - C: Wednesday
  - D: Thursday
  - E: Friday

- **n1 = Slot**
  - 1: 09.00 - 10.50
  - 2: 11.20 - 12.30
  - 3: 14.00 - 15.20
  - 4: 15.50 - 17.30

- **n2 = Track**
  - 1: Europa Hallen (ESE)
  - 2: Musiksalen (Oral)
  - 3: Radiosalen (Oral)
  - 4: The Little Theatre (Oral)
  - 5: West Hall - North (Poster)
  - 6: West Hall - South (Poster)

**Session: A21 - Oral**

**Keynotes**

- **Acquiring and implementing phonetic knowledge** .......... K-3
  - Pols L C W
  - University of Amsterdam, The Netherlands
  - For more information: http://www.fon.hum.uva.nl/pols/

- **Mobile Future** ........................................... K-7
  - Neuvo Y
  - Nokia Mobile Phones, Finland

- **How Visual Co-Presence and Joint Attention Shape Speaking** K-11
  - Brennan S E
  - Department of Psychology, State University of New York, USA

**Session: A31 - Oral**

**ESE1 - What do Industry and Universities Expect from Each Other?: Panel discussion**

- **Whither Speech Technology? - A Twenty-First Century Perspective** 3
  - Greenberg S
  - International Computer Science Institute, USA
  - For more information: http://www.icis.berkeley.edu/~steveng/PASTA

- **3G Mobile Networks and Mobile Internet as a Promotor for New Applications - Challenges to Industry and Universities** 7
  - Dobler S, Hermansson H, Minde T-B
  - Ericsson Research - Audio Visual Technology, Sweden

- **Universities and Industry: Marriage or Co-operation between Independent Partners?** 11
  - Niiniluoto I
  - University of Helsinki, Finland

- **Considerations on What Industry Expects from Universities** . 13
  - Neuvo Y
  - Nokia Mobile Phones, Finland

- **A Perspective on Industry/University Relationships in the US** 15
  - Strong G
  - Information Technology Research Program, NSF, USA

- **ELRA Contribution to Bridge the Gap Between Industry and Academia** 17
  - Choukri K
  - ELRA/ELDA, Paris, France

**Session: A32 - Oral**

**Linguistic Modelling: Language Model Compression**

- **Combining word- and class-based language models: A comparative study in several languages using automatic and manual word-clustering techniques** 21
  - IBM Voice Systems, European Speech Research

- **Multi-Class Composite N-gram Language Model Using Multiple Word Clusters and Word Successions** 25
  - Isogai S’, Shirai K’, Yamamoto H’, Sagisaka Y’
  - 1Waseda University, Japan, 2ATR Spoken Language Translation Research Laboratories, Japan

- **Statistical Language Model Based On a Hierarchical Approach: MCnV** 29
  - Zitouni I, Smaili K, Haton J-P
  - LORIA, France

- **Quantization-based Language Model Compression** 33
  - Whittaker E, Raj B
  - Compaq Cambridge Research Laboratory, USA

**Session: A33 - Oral**

**Speech Production: Voice Source**

- **Relations between vocal registers in voice breaks** 39
  - Bloothooft G, van Wijck M, Pabon P
  - Utrecht University, The Netherlands

- **A Quasi-One-Dimensional Model of Aerodynamic and Acoustic Flow in the Time-Varying Vocal Tract: Source and Excitation Mechanisms** 43
  - Ramsay G
  - Universite Libre de Bruxelles, Belgium
Spectral correlates of voice open quotient and glottal flow asymmetry: theory, limits and experimental data .......................... 47
Henrich N1, d'Alessandro C2, Doval B2
1LAM/LIMSI, France, 2LIMSI, France
For more information: http://www.limsi.fr/Individu/henrich/Eurospeech01.html

One-delayed-mass model for efficient synthesis of glottal flow .......................... 51
Avanzini F1, Paavo A2, Karjalainen M2
1Università di Padova, Italy, 2Helsinki University of Technology, Finland
For more information: http://www.dei.unipd.it/~avanzini/

Session: A34 - Oral

Modeling Pronunciation Variation Using Context-Dependent Weighting and B/S Refined Acoustic Modeling .......................... 57
Zheng F1, Song Z2, Fung P1, Byrne W2
1Tsinghua University, P. R. China, 2University of Science and Technology, Hong Kong, 3The Johns Hopkins University, USA

Learning Units for Domain-Independent Out-of-Vocabulary Word Modelling .......................... 61
Bazzi L, Glass J
MIT Laboratory for Computer Science, USA

Pronunciation Variant Analysis using Speaking Style Parallel Corpus .......................... 65
Nakajima H1, Hirano F1, Sagiaka T2, Shirai K2
1ATR Spoken Language Translation Research Laboratories, Japan, 2Waseda University, Japan

Speech recognition for huge vocabularies by using optimized subword units .......................... 69
Kneissler J, Klakow D
Philips Forschungsstaboratori, Germany

Session: A35 - Poster

Phonetics and Phonology: Prosody and Others

Factors affecting schwa-insertion in final consonant clusters in Standard Dutch .......................... 75
Swerts M, Kloots H, Gillis S, De Schutter G
Antwerp University, Belgium

Vowel Height is Intimately Associated with Stress Accent in Spontaneous American English Discourse .......................... 79
Hitchcock L, Greenberg S
International Computer Science Institute, USA
For more information: http://www.isci.berkeley.edu/~steveng/prosody

Finite State prosodic analysis of African corpus resources .......................... 83
Gibbon D
Universität Bielefeld, Germany

Acoustic Correlates of Emotion Dimensions in View of Speech Synthesis .......................... 87
Schröder M1, Cowie R1, Douglas-Cowie E2, Westerdijk M1, Gielen S1
1DFKE, Saarbrücken, Germany, 2Queen's University, Belfast, Northern Ireland, 3University of Nijmegen, The Netherlands

Measuring Pitch Range .......................... 91
Ouden, den H, Terken J
Technische Universität Eindhoven, The Netherlands

Measuring speech rhythm .......................... 95
Gibbon D, Gut U
Universität Bielefeld, Germany

Tonal alignment, scaling and slope in Italian question and statement tunes .......................... 99
D'Imperio M
LORIA - University of Nancy, France

Pragmatic temporal voice range profile as a tool in the research of speech styles .......................... 103
Ivonen A K
University of Helsinki, Finland

Model Based Stress Decision Method .......................... 107
Kim W, Kim T, Ahn S, Ko H
Korea University, Korea

Reduction of alternative pronunciations in the Norwegian computational lexicon NorKomPLeS .......................... 111
Nordgård T, Foldvik A K
Norwegian University of Science and Technology, Norway

The Role of Duration as a Correlate of Accent in Lekeitio Basque .......................... 115
Elordieta G1, Haalde J1
1University of the Basque Country, Spain, 2University of Illinois, USA

Word Final Aspiration as a Phrase Boundary Cue: Data from Spontaneous Swedish Discourse .......................... 119
Johansson V, Horne M, Strömquist S
Lund University, Sweden

Study and Auto-Detection of Stress Based on Tonal Pitch Range in Mandarin .......................... 123
Shin X, Xu B
Institute of Automation, Chinese Academy of Sciences, P. R. China

Classifying emotions in speech: a comparison of methods .......................... 127
Amir N1, Kerret O2, Karlini D2
1Holon Academic Institute of Technology, Israel, 2Academic College of Tel Aviv Yaffo, Israel

Session: A36a - Poster

Speech Perception: First and Second Language Learning

Development of vowel quantity perception in late childhood .......................... 133
Behne D M1, Czigler P E2, Sullivan K P H2
1Norwegian University of Science and Technology, Norway, 2Umeå University, Sweden

A study on the production-perception link of English vowels produced by native and non-native speakers .......................... 137
Yang B
Dongeui University, Korea

Japanese Can be Aware of Syllables and Morae: Evidence from Japanese-English Bilingual Children .......................... 141
Otake T, Yamaguchi Y
Dokkyo University, Japan
Neural Processes Underlying Perceptual Learning of a Difficult Second Language Phonetic Contrast .................................................. 145
Callan D, Tajima K, Callan A, Akahane-Yamada R, Masaki S
ATR International, Japan

Human Language Identification with Reduced Segmental Information: Comparison between Monolinguals and Bilinguals
Komatsu M1, Mori K1, Arai T1, Murahara Y2
1Sophia University, Japan / University of Alberta, Canada, 2Sophia University, Japan
For more information: http://www.splab.ee.sophia.ac.jp/

Session: A36b - Poster

Speech Perception: Miscellaneous - I
Coarticulatory effects in perception .......................... 155
Fernández S, Feijoo S
Universidad de Santiago de Compostela, Spain

A Case for Multi-Resolution Auditory Scene Analysis ....... 159
Harding S, Meyer K
Keele University, UK

Perceptual Identification and Normalization of Synthesized French Vowels from Birth to Adulthood ......... 163
Ménard L1, Schwartz J-L1, Boë L-J1, Kandel S2, Vallée N1
1Université Stendhal, France, 2Université Joseph-Fourier, France

Perceptual Categorization of Maximal Vowel Spaces from Birth to Adulthood Simulated by an Articulatory Model ...... 167
Ménard L, Boë L-J
Université Stendhal, France

A study on speech over the telephone and aging .......... 171
Eskenazi M, Black A
Carnegie Mellon University, USA

On the Perception of Voicing for Plosives in Noise .......... 175
Chen M, Alwan A
University of California at Los Angeles, USA

Predicting Visual Consonant Perception from Physical Measures .......................................................... 179
Jang J, Alwan A1, Auer E2, Bernstein L2
1University of California at Los Angeles, USA, 2House Ear Institute, Los Angeles, USA

Session: A41 - Oral & Poster

ESE2 - Noise Robust Recognition: Front-end and Compensation Algorithms

Session Introduction & Aurora Status (oral presentation, no proceedings paper) .......................... 184
Pearce D
Motorola Labs, UK

Aurora 2 Database Training and Test Sets Description (oral presentation, no proceedings paper) ............... 184
Hirsch H-G
University of Applied Sciences Niederrhein, Germany

Noise Robust Feature Extraction for ASR using the Aurora 2 Database .................................................. 185
Zhu Q, Iseli M, Cui X, Alwan A
University of California, Los Angeles, USA

Investigations into Tandem Acoustic Modeling for the Aurora Task ......................................................... 189
Ellis D P W, Reyes Gomez M J
Columbia University, USA

Recognition Performance of the Siemens Front-end with and without Frame Dropping on the Aurora 2 Database .... 193
Andrassy B1, Vlaj D2, Beagueant C2
1Siemens, München, Germany, 2University of Maribor, Slovenia

A Multiconditional Robust Front-End Feature Extraction with a Noise Reduction Procedure Based on Improved Spectral Subtraction Algorithm ...................................................... 197
Kotnik B, Kacic Z, Horvat B
University of Maribor, Slovenia

Feature Vector Selection to Improve ASR Robustness in Noisy Conditions ............................................. 201
de Veth J1, Mauuary L1, Noe B1, de Wet P2, Sienel J1, Boves L1, Jouvet D2
1University of Nijmegen, The Netherlands, 2France Télécom R&D, France, 3Alcatel SEL, Germany

Comparison of Spectral Derivative Parameters for Robust Speech Recognition ........................................... 205
Macho D, Nadeu C
TALP Research Center - UPC, Spain

Robust Digit Recognition in Noise: An Evaluation Using the AURORA Corpus ............................................. 209
Yapanel U, Hansen J H L, Sarikaya R, Pellom B
Univ. of Colorado Boulder, USA
For more information: http://cslr.colorado.edu

Robust ASR Based On Clean Speech Models: An Evaluation of Missing Data Techniques For Connected Digit Recognition in Noise ................................................................. 213
Barker J P, Cooke M, Green P
Sheffield University, UK

Evaluation of the SPLICE Algorithm on the Aurora2 Database217
Droppo J, Deng L, Acero A
Microsoft Research, USA

Model-based compensation of the additive noise for continuous speech recognition. Experiments using the AURORA II database and tasks ......................................................... 221
Segura J C, Torre A D L, Benitez M C, Peinado A M
Universidad de Granada, Spain

MAP Combination of Multi-Stream HMM or HMM/ANN Experts ...................................................... 225
Morris A C, Hagen A, Bourlard H
IDIAP, Switzerland

Second Order Statistics Spectrum Estimation Method for Robust Speech Recognition .......................................... 229
Jut B, Babic R
University of Maribor, Slovenia

Feature Extraction and Model-Based Noise Compensation for Noisy Speech Recognition evaluated on AURORA 2 Task ... 233
Yao K, Chen J, Paliwal K K, Nakamura S
ATR Spoken Language Translation Research Labs., Japan

Session: A42 - Oral

Linguistic Modelling: Language Model Adaptation
Broadcast News LM Adaptation using Contemporary Texts 239
Federico M, Bertoldi N
ITC-Ist, Italy

Topic Detection for Language Model Adaptation of Highly-Inflected Languages by Using a Fuzzy Comparison Function 243
Sepesy Maucen M, Kacic Z
University of Maribor, Slovenia

Efficient Stochastic Finite-State Networks for Language Modelling in Spoken Dialogue Systems 247
Georgiia K, Fakotakis N, Kokkinakis G
University of Patras, Greece

Language Models Conditioned on Dialog State 251
Visweswariah K, Printz H
IBM, USA

Using Information Retrieval Methods for Language Model Adaptation 255
Chen L, Gauvain J-L, Lamel L, Adda G, Adda-Decker M
CNRS-LIMSI, France

Speech Production: Articulation

Making the Tongue Model Talk: Merging MRI & EMA Measurements 261
Engwall O
KTH, Sweden
For more information: http://www.speech.kth.se/multimodal/vocaltract.html

The Relationship between Intraoral Air Pressure and Tongue/Palate Contact during the Articulation of Norwegian /t/ and /d/ 265
Moen I, Simonsen H G, Huseby M, Grue J
University of Oslo, Norway

Mechanical versus perceptual constraints as determinants of articulatory strategy 269
Elgendy A M, Pols L C W
University of Amsterdam, The Netherlands

Pre-Liquid Excrescent Schwa: What Happens when Vocalic Targets Conflict 273
Gick B, Wilson I
University of British Columbia, Canada

Exploring the Null Space of the Acoustic-to-Articulatory Inversion Using a Hypercube Codebook 277
Omni S, Laprie Y
LORIA, France

Speech Recognition and Understanding: Topic Detection and Information Retrieval

Phoneme-based Topic Spotting on the Switchboard Corpus 283
Theunissen M W, Scheffler K', du Preez J A'
1University of Stellenbosch, South Africa, 2University of Cambridge, United Kingdom

Topic Styles in IR and TDT: Effect on System Behavior ... 287
IBM TJ Watson Research Center, USA

Extracting Caller Information from Voicemail 291
Zweig G, Huang J, Padmanabhan M
IBM TJ Watson Research Center, USA

A Portability Study on Natural Language Call Steering 295
Kuo H-K J, Lee C-H
Bell Labs, Lucent Technologies, USA

Improved Spoken Document Retrieval by Exploring Extra Acoustic and Linguistic Cues 299
Chen B, Wang H-M, Lee L-S
Institute of Information Science, Academia Sinica, Taiwan, ROC

Phonetics and Phonology: Segmentals and Synthesis

Native vs Non-native Production of English Vowels in Spontaneous Speech: An Acoustic Phonetic Study 305
Tsukada K
Macquarie University, Australia
Non-english Abstract: ne1031.pdf

Is Non-Native Pronunciation Modelling Necessary? 309
Goronzy S', Sahakyan M', Wokurek W'
1Sony International (Europe), Germany, 2IMS, University of Stuttgart, Germany

Burst segmentation and evaluation of acoustic cues 313
Laprie Y, Bonnaud A
LORIA, France

The schwa in Albanian 317
Granzer T, Moosmüller S
Institute of Acoustics, Austrian Academy of Sciences, Austria

A Testbed for Developing Multilingual Phonotactic Descriptions 321
Ashby S, Carson-Berndsen J, Joue G
University College Dublin, Ireland

A Physiological Analysis of Nasals and Nasalization in Chinese 325
Fung W-N, Lau S-L
City University of Hong Kong, Hong Kong

A Component by Component Listening Test Analysis of the IBM Trainable Speech Synthesis System 329
Donovan R E
IBM TJ Watson Research Center, USA

Semantic Abnormality and its Realization in Spoken Language333
Pan S, McKeown K, Hirschberg J
1IBM TJ Watson Research Center, USA, 2Columbia University, USA, 3AT&T Labs-Research, USA

TALKING FOREIGN - Concatenative Speech Synthesis and the Language Barrier 337
Campbell N
ATR Spoken Language Translation, Japan
For more information: http://www.slt.atr.co.jp/~nick
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schwa-assimilation in Danish Synthetic Speech</td>
<td>341</td>
</tr>
<tr>
<td>Jensen C</td>
<td></td>
</tr>
<tr>
<td>University of Copenhagen, Denmark</td>
<td></td>
</tr>
<tr>
<td>For more information: <a href="http://www.cphling.dk/pers/chrijen/eurospeech2001/">http://www.cphling.dk/pers/chrijen/eurospeech2001/</a></td>
<td></td>
</tr>
<tr>
<td>Text-to-speech synthesis with arbitrary speaker's voice from average voice</td>
<td>345</td>
</tr>
<tr>
<td>Tamura M, Masuko T, Tokuda K, Kobayashi T</td>
<td></td>
</tr>
<tr>
<td>Tokyo Institute of Technology, Japan</td>
<td></td>
</tr>
<tr>
<td>Nara Institute of Science and Technology, Japan</td>
<td></td>
</tr>
<tr>
<td>High Quality Voice Conversion Based on Gaussian Mixture Model</td>
<td>349</td>
</tr>
<tr>
<td>with Dynamic Frequency Warping</td>
<td></td>
</tr>
<tr>
<td>Toda T, Saruwatari H, Shikano K</td>
<td></td>
</tr>
<tr>
<td>Nara Institute of Science and Technology, Japan</td>
<td></td>
</tr>
<tr>
<td>Evaluation of Cross-Language Voice Conversion Based on GMM and Straight</td>
<td>357</td>
</tr>
<tr>
<td>Universidad Politécnica de Madrid, Spain</td>
<td></td>
</tr>
<tr>
<td>Voice Transformations: From Speech Synthesis to Mammalian Vocalizations</td>
<td>353</td>
</tr>
<tr>
<td>Tang M, Wang C, Seneff S</td>
<td></td>
</tr>
<tr>
<td>MIT Laboratory for Computer Science, USA</td>
<td></td>
</tr>
<tr>
<td>A new Multi-Speaker Formant Synthesizer that applies Voice</td>
<td>357</td>
</tr>
<tr>
<td>Conversion Techniques</td>
<td></td>
</tr>
<tr>
<td>Universidad Politécnica de Madrid, Spain</td>
<td></td>
</tr>
<tr>
<td>Evaluation of Cross-Language Voice Conversion Based on GMM and Straight</td>
<td>361</td>
</tr>
<tr>
<td>Mashimo M, Toda T, Shikano K, Campbell N</td>
<td></td>
</tr>
<tr>
<td>Nara Institute of Science and Technology, Japan</td>
<td></td>
</tr>
<tr>
<td>Information Sciences Division, Japan</td>
<td></td>
</tr>
<tr>
<td>Ejective Reduction in Chaha is Conditioned by More Than Prosodic Position</td>
<td>365</td>
</tr>
<tr>
<td>Coulston R</td>
<td></td>
</tr>
<tr>
<td>Oregon Graduate Institute / University of California San Diego, USA</td>
<td></td>
</tr>
<tr>
<td>For more information: <a href="http://www.geocities.com/rachelcoulston">http://www.geocities.com/rachelcoulston</a></td>
<td></td>
</tr>
<tr>
<td>Speech Perception: Miscellaneous - II</td>
<td></td>
</tr>
<tr>
<td>Session: ESE2 - Noise Robust Recognition: Front-end Algorithms</td>
<td></td>
</tr>
<tr>
<td>Effects of Noise Adaptation on the Perception of Voiced Plosives in Isolated Syllables</td>
<td>371</td>
</tr>
<tr>
<td>Ainsworth W, Cervera T</td>
<td></td>
</tr>
<tr>
<td>Keele University, UK, University of Valencia, Spain</td>
<td></td>
</tr>
<tr>
<td>On Differential Linen of Word-based Local Speech Rate Variation in Japanese Expressed by Duration Ratio</td>
<td>375</td>
</tr>
<tr>
<td>Hiroshage M, Araki K, Tochinai K</td>
<td></td>
</tr>
<tr>
<td>Hokkaido University, Sapporo, Japan</td>
<td></td>
</tr>
<tr>
<td>A Multidimensional Scaling Study of Fricatives; a Comparison of Perceptual and Physical Dimensions</td>
<td>379</td>
</tr>
<tr>
<td>Tokuma W</td>
<td></td>
</tr>
<tr>
<td>Seijo University, Japan</td>
<td></td>
</tr>
<tr>
<td>Reconstructing Dialogue History</td>
<td>383</td>
</tr>
<tr>
<td>Swerts M, Kramer E</td>
<td></td>
</tr>
<tr>
<td>IPO, The Netherlands / CNTS, Belgium, IPO, The Netherlands</td>
<td></td>
</tr>
<tr>
<td>Timing and Interaction of Visual Cues for Prominence in Audiovisual Speech Perception</td>
<td>387</td>
</tr>
<tr>
<td>House D, Beskow J, Granström B</td>
<td></td>
</tr>
<tr>
<td>KTH, Sweden</td>
<td></td>
</tr>
<tr>
<td>Modelling the Perceptual Identification of Japanese Consonants from LPC Cepstral Distances</td>
<td>391</td>
</tr>
<tr>
<td>Komatsu M, Tokuma S, Tokuma W, Arai T</td>
<td></td>
</tr>
<tr>
<td>Sophia University, Japan / University of Alberta, Canada, Sagami Women's University, Japan, Sophia University, Japan</td>
<td></td>
</tr>
<tr>
<td>For more information: <a href="http://www.splab.ee.sophia.ac.jp/">http://www.splab.ee.sophia.ac.jp/</a></td>
<td></td>
</tr>
<tr>
<td>Auditory-Visual Perception of Lexical Tone</td>
<td>395</td>
</tr>
<tr>
<td>Barnhaim D, Ciocca V, Stokes S</td>
<td></td>
</tr>
<tr>
<td>University of Western Sydney, Australia, University of Hong Kong, Hong Kong</td>
<td></td>
</tr>
<tr>
<td>Syllable prominence: A matter of vocal effort, phonetic distinctness and top-down processing</td>
<td>399</td>
</tr>
<tr>
<td>Eriksson A, Thuanberg G C, Traummüller H</td>
<td></td>
</tr>
<tr>
<td>Stockholm University, Sweden</td>
<td></td>
</tr>
<tr>
<td>Perceived Prominence in Terms of a Linguistically Motivated</td>
<td>403</td>
</tr>
<tr>
<td>Quantitative Intonation Model</td>
<td></td>
</tr>
<tr>
<td>Mixdorff H, Widera C</td>
<td></td>
</tr>
<tr>
<td>Berlin University of Applied Sciences, Germany, University of Bonn, Germany</td>
<td></td>
</tr>
<tr>
<td>Perception of Coda Voicing from Properties of the Onset and Nucleus of 'led' and 'let'</td>
<td>407</td>
</tr>
<tr>
<td>Hawkins S, Nguyen N</td>
<td></td>
</tr>
<tr>
<td>University of Cambridge, UK, Université de Provence, France</td>
<td></td>
</tr>
<tr>
<td>For more information: <a href="http://www.ill.pt/uni_uk-fv/gnuen/let-led.html">http://www.ill.pt/uni_uk-fv/gnuen/let-led.html</a></td>
<td></td>
</tr>
<tr>
<td>Auditory Filter Bank Design Using Masking Curves</td>
<td>411</td>
</tr>
<tr>
<td>Lin L L, Ambikairajah E, Holmes W H</td>
<td></td>
</tr>
<tr>
<td>The University of New South Wales, Australia</td>
<td></td>
</tr>
<tr>
<td>A New Feature Driven Cochlear Implant Speech Processing Strategy</td>
<td>415</td>
</tr>
<tr>
<td>Dashitioner E, Kitazawa S, Kitamura T</td>
<td></td>
</tr>
<tr>
<td>Shizuoka University, Japan</td>
<td></td>
</tr>
</tbody>
</table>

Session: ESE2 - Noise Robust Recognition: Front-end Algorithms

Acoustic Feature Compensation Based on Decomposition of Speech and Noise for ASR in Noisy Environments | 421 |
Kim H K, Rose R C, Kang H G | AT&T Labs-Research, USA |

A Robust Front-End Algorithm for Distributed Speech Recognition | 425 |
Cheng Y M, Macho D, Wei Y, Ealey D, Kelleher H, Pearce D, Kushliner W, Ramabadran T | Motorola Labs, USA |

Robust ASR front-end using spectral-based and discriminant features: experiments on the Aurora tasks | 429 |
Benitez C, Burgst L, Chen B, Dupont S, Garudadri H, Hermansky H, Jain P, Kazajekar S, Sivadas S | International Computer Science Institute, USA, Oregon Graduate Institute Of Science and Technology, USA, Qualcomm Inc., USA |

Noise Reduction for Noise Robust Feature Extraction for Distributed Speech Recognition | 433 |
Noel B, Sienel F, Jouvet D, Mauuary L, de Veth F, Boves L, de Wet F | Alcatel SEL, Stuttgart, Germany, France Telecom R&D, DIH/IPS, University of Nijmegen, The Netherlands |
Harmonic tunnelling: tracking non-stationary noises during speech .......................................................... 437
Ealey D, Kelleher H, Pearce D
Motorola Labs, UK

Session: B12 - Oral

Linguistic Modelling: Semantic Modelling

Resource-Limited Sentence Boundary Detection .................. 443
Carter D, Grandsen I
Speech Machines, UK

Metrics for Measuring Domain Independence of Semantic Classes .......................................................... 447
Bell Labs, Lucent Technologies, USA

Context-dependent Probabilistic Hierarchical Sub-lexical Modelling Using Finite State Transducers ............... 451
Mou X, Benefi S, Zue V
MIT Laboratory for Computer Science, USA

Data-Driven Semantic Inference for Unconstrained Desktop Command and Control ........................................ 455
Bellegarda J, Silverman K
Apple Computer, USA

Information Extraction via Heuristics for a Movie Showtime Query System ..................................................... 459
Jansche M
The Ohio State University, USA

Session: B13 - Oral

Speech Perception: Recognition and Intelligibility

Recognition of (Almost) Spoken Words: Evidence from Word Play in Japanese ............................................. 465
Otake T, Cutler A
1Dokkyo University, Japan, 2Max Planck Institute for Psycholinguistics, The Netherlands

Perceptual experiments on enhanced and slowed down speech sentences for second language acquisition ........... 469
Colotte V, Laprie Y, Bonneau A
LORIA, France

The Relation Between Speech Intelligibility and the Complex Modulation Spectrum ......................................... 473
Greenberg S, Ari Y
1International Computer Science Institute, USA, 2Sophia University, Japan
For more information: http://www.icsi.berkeley.edu/~steveng

Envelope Information in Speech Processing: Acoustic-Phonetic Analysis vs. Auditory Figure-Ground Segregation ........ 477
Crouzet O, Ainsworth W A
Keele University, UK

A comparison between human vowel normalization strategies and acoustic vowel transformation techniques ........ 481
Adank P, van Hout R, Smits R
University of Nijmegen, The Netherlands

Session: B14 - Oral

Speech Recognition and Understanding: LVCSR - I

On Large Vocabulary Continuous Speech Recognition of Highly Inflectional Language - Czech ................................ 487
1University of West Bohemia in Pilzen, Czech Republic, 2MFF UK, Czech Republic, 3Johns Hopkins University, USA

Towards Automatic Transcription of Spontaneous Presentation91
Shinozaki T, Hori C, Furui S
Tokyo Institute of Technology, Japan
For more information: http://www.furui.cs.titech.ac.jp

A Real-Time Japanese Broadcast News Closed-Captioning System ............................................................. 495
1Bell Labs - Lucent Technologies, USA, 2NHK - Science and Technical Research Laboratories, Japan
For more information: http://www.bell-labs.com/user/siohan

Investigations on Conversational Speech Recognition ......... 499
Beyerlein P, Aubert X, Harris M, Meyer C, Schramm H
Philips Research Laboratories Aachen, Germany

Recent Advances in Speech Recognition System for IBM DARPA Communicator ............................................ 503
Gao Y, Erdogan H, Li Y, Goel V, Picheny M
IBM T. J. Watson Research Center, USA

Session: B15 - Poster

Speech Synthesis: Systems and Prosody

Festival Speaks Italian! ............................................. 509
Cosi P, Tesser F, Greetter R, Avanesi C, Macon M
1Istituto di Fonetica e Dialettologia - Consiglio Nazionale delle Ricerche, Italy, 2Istituto Trentino di Cultura - Istituto per la Ricerca Scientifica e Tecnologica, Italy, 3Oregon Graduate Institute for Science and Technology, USA

Multilingual TTS for Computer Telephony: The Aculab Approach ............................................................ 513
Aculab plc, UK
For more information: http://www.aculab.com

A Flexible Multilingual TTS Development and Speech Research Tool .......................................................... 517
Kiss G, Németh G, Olasy G, Gordes G
1Budapest University of Technology and Economics, Hungary, 2Hungarian Academy of Sciences, Hungary
For more information: http://speechlab.ttt.bme.hu

Speech Synthesis Development Made Easy: The Bonn Open Synthesis System ............................................. 521
1IPPO, Center for User-System Interaction, the Netherlands, 2IKP, University of Bonn, Germany

Automatic Prosody Generation - a Model for Hungarian ...... 525
Olasy G, Németh G, Olasi P
1Hungarian Academy of Sciences, Hungary, 2Budapest University of Technology and Economics, Hungary
For more information: http://speechlab.ttt.bme.hu
Evaluation of PROS-3 for the assignment of prosodic structure, compared to assignment by human experts .......................... 529
van Herwijnen O M, Terken J M B
Technische Universiteit Eindhoven, The Netherlands

Stochastic F0 Contour Model Based on the Clustering of F0 Shapes of a Syntactic Unit ........................................ 533
Yamashita Y, Ishida T
Ritsumeikan University, Japan

Intonational Phrase Break Prediction Using Decision Tree and N-Gram Model ........................................ 537
Sun X1, Applebaum T H1
1Northwestern University, USA, 2Panasonic Speech Technology Laboratory, USA

Synthesizing Intonation of Standard Arabic Language . . . 541
Zaki A1, Rajouani A2, Najim M1
1Equipe Signal et Image, ENSEIRB, France, 2Laboratoire d’Electronique et Etudes des Systèmes Automatiques Rabat, Morocco
For more information: http://www.its.u-bordeaux.fr/zaki/arabic-synthesis-demo.htm
Non-english Abstract: ne1023.pdf

Invariance of Relative F0 Change Field of Chinese Disyllabic Words ........................................ 545
Xu D, Mori H, Kasuya H
Okayama University, Japan

Accent Label Prediction by Time Delay Neural Networks Using Gating Clusters ........................................ 549
Mueller A F1, Hoffmann R2
1Siemens, Germany, 2Dresden University of Technology, Germany

Transformation-Based Learning of Danish Stress Assignment 553
Henrichsen P J
Copenhagen Business School, Denmark

On the Prosody of German Telephone Numbers ........... 557
Baumann S, Trouvain J
University of the Saarland, Germany

Emotional Speech Synthesis: A Review .......................... 561
Schroder M
DFKI, Saarbrucken, Germany

Fun or Boring? A Web-based Evaluation of Expressive Synthesis for Children ........................................ 565
Gustafson K, House D
KTH, Sweden

Session: B16 - Poster

Speach Recognition and Understanding: Articulatory and Perceptual Approaches to ASR

Sub-Band Based Additive Noise Removal for Robust Speech Recognition ........................................ 571
Chen J1, Palival K K1, Nakamura S1
1ATR Spoken Language Translation Research Laboratories, Japan, 2Griffith University, Australia

Development of an Asynchronous Multi-band System for Continuous Speech Recognition .......................... 575
Tam Y-C, Maik B
The Hong Kong University of Science and Technology, Hong Kong

A Multi-Band Approach Based on the Probabilistic Union Model and Frequency-Filtering Features for Robust Speech Recognition ........................................ 579
Jancovic P, Ming J
Queen’s University of Belfast, UK

Split-band Perceptual Harmonic Cepstral Coefficients as Acoustic Features for Speech Recognition ............... 583
Gu L, Rose K
University of California, Santa Barbara, USA

Error Correcting Posterior Combination for Robust Multi-Band Speech Recognition .......................... 587
Hagen A, Bourlard H
IDIAP, Martigny, Switzerland

Robust Parameters for Speech Recognition Based on Subband Spectral Centroid Histograms ......................... 591
Gagic B1, Palival K1
1Norwegian University of Science and Technology, Norway, 2Griffith University, Australia

Pseudo-Articulatory Representations and the Recognition of Syllable Patterns in Speech .......................... 595
Edmondson W, Zhang L
University of Birmingham, UK

ASR - Articulatory Speech Recognition .......................... 599
Frankel J, King S
University of Edinburgh, UK
For more information: http://www.cstr.ed.ac.uk

Efficient Decoding Strategy for Conversational Speech Recognition Using State-Space Models for Vocal-Tract-Resonance Dynamics ........................................ 603
Ma J1, Deng L2
1BNB Technologies, USA, 2Microsoft Company, USA

HMM2- Extraction of Formant Structures and their Use for Robust ASR ........................................ 607
Weber K1, Bengio S2, Bourlard H1
1IDIAP, Martigny / EPFL, Lausanne, Switzerland, 2IDIAP, Martigny, Switzerland

Auditory model based speech recognition in noisy environment611
Yu X, Wan W, Lan D
The Hong Kong Polytechnical University, Hong Kong

Forward Masking for Increased Robustness in Automatic Speech Recognition .......................... 615
Wendt S, Fink G A, Kummert F
Universitaet Bielefeld, Germany
Non-english Abstract: ne1034.pdf

An Auditory System-Based Feature for Robust Speech Recognition ........................................ 619
Li Q, Soong F K, Olivier S
Bell Labs, Lucent Technologies, USA

Session: B21 - Oral

ESE2 - Noise Robust Recognition: Robust systems - What helps?

Experiments with the Philips continuous ASR system on the AURORA noisy digits database .......................... 625
Lieb M, Fischer A
Philips Research Labs Aachen, Germany
Robust Digit Recognition in Noisy Environments: the IBM Aurora 2 System .................. 629
Saon G, Huerta J, Jan E-E
IBM T.J. Watson Research Center, USA

Evaluating the Aurora Connected Digit Recognition Task -- A Bell Labs Approach .................. 633
Bell Labs, Lucent Technologies, USA

Improved Maximum Mutual Information Estimation Training of Continuous Density HMMs .................. 679
Zheng J, Butzberger J,Franco H, Stolcke A
SRI International, USA

Maximum-Likelihood Training of a Bipartite Acoustic Model for Speech Recognition .................. 683
Peronnin F, Kuhn R, Nguyen P,Junqua J-C
Panasonic Speech Technology Laboratory, USA

Analysis of the Root-Cepstrum for Acoustic Modeling and Fast Decoding in Speech Recognition .................. 687
Sarikaya R, Hansen J
RSPL-CLSER, University of Colorado-Boulder, USA

Session: B22 - Oral

Phonetics and Phonology: Segmentals

Liaison and schwa deletion in French: an effect of lexical frequency and competition? .................. 639
Fougeron C, Goldman J-P, Frauenfelder U H
Univ. of Geneva, Switzerland

An Acoustical Analysis of the Vowels in Beijing Mandarin .................. 643
Zoe E1, Lee W-S2
1City University of Hong Kong, Hong Kong, 2University of Hong Kong, Hong Kong

Discriminant analysis of nasal vs. oral vowels in French: comparison between different parametric representations .................. 647
Delvaux V, Soquet A
Université Libre de Bruxelles, Belgium

Whispery voiced nasal stops in Rwanda .................. 651
Demolín D, Delvaux V
Université Libre de Bruxelles, Belgium

Session: B23 - Oral

Speech Production: Prosody - I

Prominence correlates. A study of Swedish .................. 657
Fant G1, Kruckenberg A2, Liljencrate J, Botinis A2
1KTH, Sweden, 2University of Skövde, Sweden

Quantitative Analysis of the Effects of Emphasis Upon Prosodic Features of Speech .................. 661
Ohno S1, Fujisaki H2
1Tokyo University of Technology, Japan, 2Science University of Tokyo, Japan

Towards a Model of Target Oriented Production of Prosody .................. 665
Dogil G, Mohus B
University of Stuttgart, Germany

Prosody Control for Speaking and Singing Styles .................. 669
Shih C, Kocauski G P
Bell Laboratories, Lucent Technologies, USA
For more information: http://www.bell-labs.com/project/tts/stem.html

Session: B24 - Oral

Speech Recognition and Understanding: Acoustic Modelling - I

A Mixture of Gaussians Front End for Speech Recognition .................. 675
Stuttle M, Gales M J F
Cambridge University, UK

Speech Recognition of Broadcast Sports News .................. 709
Matsui A, Segi H, Kobayashi A, Ima T, Ando A
NHK(Japan Broadcasting Corp.), Japan

The Study Of The Effect Of Training Set On Statistical Language Modeling .................. 721
Shen X, Xu B
Institute of Automation, Chinese Academy of Sciences, P.R. China

Session: B25 - Poster

Linguistic Modelling: Language Models - I

Structured language model for class identification of out-of-vocabulary words arising from multiple word-classes .................. 693
Onishi S, Yamamoto H, Sagisaka Y
ATR Spoken Language Translation Research Laboratories, Japan

New Language Models Using Phrase Structures Extracted From Parse Trees .................. 697
Jitsuhiro T1, Yamamoto H2, Yamada S, Sagisaka Y
1ATR Spoken Language Translation Research Laboratories, Japan, 2NTT Communication Science Laboratories, Japan
Non-english Abstract: ne1114.pdf

A Structured Statistical Language Model conditioned by Arbitrarily Abstracted Grammatical Categories based on GLR Parsing .................. 705
Akiba T, Itoh K
National Institute of Advanced Industrial Science and Technology, Japan

Speech Recognition of Broadcast Sports News .................. 717
Kim W, Khudanpur S, Wu J
The Johns Hopkins University, USA

The Study Of The Effect Of Training Set On Statistical Language Modeling .................. 725
Esteve Y1, Bechet F1, Nast A2, De Mori R1
1LIA - University of Avignon, France, 2LIM - University of Aix-Marseille II, France
A Baseline Method for Compiling Typed Unification Grammars into Context Free Language Models ......................... 729
Rayner M1, Dowling J1, Hockey B A1
1Netdecisions, UK, 2RIACS, USA

Comparison of Width-wise and Length-wise Language Model Compression ................................................... 733
Whittaker E, Raj B
Comp Cambridgeshire Research Laboratory, USA

Large Vocabulary Statistical Language Modeling for Continuous Speech Recognition in Finnish .................... 737
Sivola V, Kurimo M, Lagus K
Helsinki University of Technology, Finland
For more information: http://www.cis.hut.fi/mikkok/speech.shtml
Non-english Abstract: ne1774.html

A New Technique Based on Augmented Language Models to Improve the Performance of Spoken Dialogue Systems .... 741
López-Cózar R1, Milone D H1
1Granada University, Spain, 2Universidad Nacional de Entre Ríos, Argentine

Session: B26 - Poster

Speaker Recognition: Identification, Verification and Tracking, Speech Recognition and Understanding: Language Identification
The Influence of Vocal Effort on Human Speaker Identification#47
Brungart D S1, Scott K R1, Simpson B D2
1Air Force Research Laboratory, USA, 2Veridian, USA

Improving Speaker Recognition Using Phonetically Structured Gaussian Mixture Models ............................ 751
Faltlhauser R, Ruske G
Technische Universität Muenchen, Germany

Information Fusion for Robust Speaker Verification ............. 755
Sanderson C, Paliwal K
Griffith University, Australia

A Robust Speaker Verification System against Imposture Using an HMM-based Speech Synthesis System .......... 759
Satoh T1, Masuko T1, Kobayashi T1, Tokuda K2
1Tokyo Institute of Technology, Japan, 2Nagoya Institute of Technology, Japan

Sequential Decisions for Faster and More Flexible Verification#63
Surendran A
Bell Labs, Lucent Technologies, USA

Background Learning of Speaker Voices for Text-independent Speaker Identification ................................. 767
Tsai W-H1, Chu Y C1, Huang C-S1, Chang W-W1
1Philips Research East Asia-Taipei / National Chiao Tung University, Taiwan, ROC, 2Philips Research East Asia-Taipei, Taiwan, ROC, 3National Chiao Tung University, Taiwan, ROC

Explicit Exploitation of Stochastic Characteristics of Test Utterance for Text-independent Speaker Identification .... 771
Tsai W-H1, Chang W-W1, Huang C-S1
1Philips Research East Asia-Taipei / National Chiao Tung University, Taiwan, ROC, 2Philips Research East Asia-Taipei, Taiwan, ROC

Improvement of Speaker Verification for Thai Language ........ 775
Wutiwiwatchai C, Achariyakulkorn V, Kasuniya S
National Electronics and Computer Technology Center, Thailand

Speaker Identification for Car Infotainment Applications ........ 779
Rodriguez-Saeta J1, Koechling C2, Hernandez J1
1UPC, Spain, 2Robert Bosch, Germany
Non-english Abstract: ne1780.pdf

A System for Text Dependent Speaker Verification - Field Trial Evaluation and Simulation Results .................. 783
Schalk H1, Reininger H1, Euler S1
1Universitaet Frankfurt, Germany, 2ATIP, Germany, 3Robert Bosch, Germany

Speaker Recognition in a Multi-Speaker Environment ........ 787
Martin A, Przybecki M
National Institute of Standards and Technology, USA

Preliminary experiments on language identification using broadcast news recordings ............................. 799
Benarousse L, Geoffrois E
DGA/CTA/GIP, France

Multi-Stream Statistical N-Gram Modeling With Application To Automatic Language Identification .................... 803
Kirchoff K, Parandekar S
University of Washington, USA

Session: B31 - Demonstrations

ESE3 - Imagination 2001 .................................................. 808

Session: B32 - Oral

Phonetics and Phonology: Pronunciation and Timing

Up to what level can acoustical and textual features predict prominence ............................................. 811
Streefkerk B M1, Pols L C W1, ten Bosch L F M2
1Univ. of Amsterdam, the Netherlands, 2L & H, Belgium

Linguistic Factors Affecting Timing in Korean with Application to Speech Synthesis ................................. 815
Chung H, Hackvale M
University College London, UK
For more information: http://www.phon.ucl.ac.uk/home/hchung/home.htm

Measuring Rhythmic Deviation in Second Language Speech 819
Schaeffler F
University of Munich, Germany
Good timing: Place-dependent voice onset time in ejective stops  823
Maddieson I
University of California, Berkeley, USA

Session: B33 - Oral

Speech Synthesis: Concatenation - I

Design of an optimal continuous speech database for text-to-
speech synthesis considered as a Set Covering Problem . . . . . . 829
Francois H, Boeffard O
IRISA, Université Rennes 1, ENSSAT, France

Use of Clustering Information for Coarticulation Compensation
in Speech Synthesis by Word Concatenation ......................... 833
Vosnidis C, Digalakvin T
Technical University of Crete, Greece

Reducing spectral mismatches in concatenative speech synthesis
via systematic database enrichment. .............................. 837
Founda M, Tambouratzis G, Chalamanaris A, Carayannis G
Institute for Language and Speech Processing, Greece

Hansori 2001 - Corpus-based Implementation of the Korean
Hansori Text-to-Speech Synthesizer ......................... 841
Ferencz A, Choi S-W, Song H-E, Koo M-W
Korea Telecom R & D Group, Korea

Session: B34 - Oral

Speech Recognition and Understanding: LVCSR - II

Time and Memory Efficient Viterbi Decoding for LVCSR using a
Precompiled Search Network .................................. 847
Willett D, McDermott E, Minami Y, Katagiri S
NTT Communication Science Laboratories, Japan

A New Verification-Based Fast Match Approach to Large
Vocabulary Speech Recognition .................................. 851
Liu F, Affi M, Jiang H, Siohan O
Bell Laboratories, Lucent Technologies, USA

A Fast Calculation Method in LVCSR by Time-Skipping and
Clustering of Probability Density Distributions .................. 855
Nakagawa S, Horibe Y
Toyohashi University of Technology, Japan

Speech Recognition of Japanese News Commentary ........... 859
Homma S, Kobayashi A, Sato S, Imai T, Ando A
NHK (Japan Broadcasting Corp.), Japan

Session: B35 - Poster

Speech Recognition and Understanding: Noise
Robustness - I

A comparison of LPC and FFT-based acoustic features for noise
robust ASR .................................................. 865
de Wet F, Cranen B, de Veth J, Boves L
University of Nijmegen, The Netherlands

Unsupervised Noisy Environment Adaptation Algorithm Using
MLLR and Speaker Selection .................................. 869
Yamada M1, Baba A1, Yoshizawa S1, Mera Y1, Lee A1, Saruwatari H1,
Shikano K1
1Nara Institute of Science and Technology, Japan, 2Laboratories of
Image Information Science and Technology, Japan

Applying Parallel Model Compensation with Mel-Frequency
Discrete Wavelet Coefficients for Noise-Robust Speech
Recognition .................................................. 873
Tufekci Z, Gowdy I, Garby S, Patterson E
Clemson University, USA

Linear Interpolation of Cepstral Variance for Noisy Speech
Recognition .................................................. 877
Hwang T-H1, Yoo K-H2, Wang H-C2
1Industrial Technology Research Institute, Taiwan, ROC, 2National
Tsing Hua University, Taiwan, ROC

Evaluation of a Generalized Dynamic Cepstrum in Distant
Speech Recognition ........................................ 881
Matsumoto H, Shimizu A, Yamamoto K
Shinshu University, Japan

Robust Speech/Non-Speech Detection using LDA applied to
MFCC for continuous Speech Recognition ...................... 885
Martin A, Dammati G, Mauuary L
France Telecom R&D, France

Toward Noise-tolerant Acoustic Models ......................... 889
Trentin E, Gori M
Università di Siena, Italy

Noise Estimation Without Explicit Speech, Non-speech Detection:
a Comparison of Mean, Modal and Median Based Approaches 893
Evans N W D, Mason J S
University of Wales, UK
For more information: http://gallle.swan.ac.uk

Evaluation of Front-End Features and Noise Compensation
Methods for Robust Mandarin Speech Recognition ........... 897
Chengalvarayan R
Lucent Technologies, USA

ALGONQUIN: Iterating Laplace's Method to Remove Multiple
Types of Acoustic Distortion for Robust Speech Recognition 901
Frey B J1, Deng L2, Acero A3, Kristjansson T
1University of Toronto, Canada, 2Microsoft Research, USA,
3University of Waterloo, Canada
For more information: http://www.cs.toronto.edu/~frey

Robust Speech Recognition in Noise: An Evaluation using the
SPINE Corpus ............................................... 905
Hansen J H L, Sarikaya R, Yapanel U, Pellom B
Univ. of Colorado Boulder, USA

Session: B36 - Poster

Speech Production: Prosody - II

Automated modeling of Chinese intonation in continuous speech 1
Kochanski G, Shih C
Bell Laboratories, Lucent Technologies, USA
For more information: http://www.bell-labs.com/project/tts/stem.html
Prediction of intonation patterns of accented words in a corpus of read Swedish news through pitch contour stylization .......................... 915
Frid J
Lund University, Sweden

The use of fundamental frequency raising as a strategy for increasing vocal intensity in soft, normal, and loud phonation...... 919
Alku P1, Vinturi J2, Vikman E1
1Helsinki University of Technology, Finland, 2Helsinki University Central Hospital, Finland, 3Oulu University, Finland

Prosodic Interactions on Segmental Durations In Greek .......................... 923
Botinis A1, Fourakis M2, Bannert R3
1University of Skövde, Sweden, 2The Ohio State University, USA, 3Umeå University, Sweden

Study on Factors Influencing Durations of Syllables in Mandarin ................................................................................ 927
Chu M1, Feng Y2
1Microsoft Research China, P. R. China, 2Institute of Acoustics, Chinese Academy of Sciences, P. R. China

A Comparative Study of Pauses in Dialogues and Read Speech031
Gustafson-Capkova S1, Megyesi B1
1Stockholm University, Sweden, 2KTH, Sweden

Detecting Japanese Local Speech Rate Deceleration in Spontaneous Conversational Speech Using a Variable Threshold
Takamaru K, Hiroshige M, Araki K, Tochinai K
Hokkaido University, Japan

Modelling Fundamental Frequency in First Post-tonic Syllables in Danish Sentences .......................................................... 939
Petersen N R
University of Copenhagen, Denmark

Non-Finality and Pre-Finality in Bari Italian Intonation: a Preliminary Account ............................................................... 943
Savino M
Politecnico di Bari, Italy

Building An Integrated Prosodic Model of German ........................................ 947
Mixdorff H1, Jokisch O2
1Berlin University of Applied Sciences, Germany, 2Dresden University of Technology, Germany
For more information: http://www.tfb-berlin.de/~mixdorff/essv98/eurospeech2001.htm

A Model of F0 Contour for Arabic Affirmative and Interrogative Sentences ...................................................................... 951
Ibrahim O A G1, El-Ramly S H1, Abdel-Kader N S2
1 Ain Shams University, Egypt, 2 Cairo University, Egypt

Variation in Final Lengthening as a Function of Topic Structure .......................... 955
Smith C L, Hogan L A
University of New Mexico, USA
For more information: http://www.unm.edu/~caroline/grant.html

Do speakers realize the prosodic structure they say they do? .... 959
van Herwijnen O M, Terken J M B
Technische Universität Eindhoven, The Netherlands

Coarticulatory Effects at Prosodic Boundaries: Some Acoustic Results ................................................................................ 963
Tabain M, Rolland G, Savariaux C
Institut de la Communication Parlée, France
For more information: http://www.shrc.mq.edu.au/~marija
Non-english Abstract: ne1455.html

Generating Duration from a Cognitively Plausible Model of Rhythm Production ................................................................. 967
Barbosa P
IEL/UNICAMP, Brazil
Non-english Abstract: ne1207.html

Session: B41 - Demonstrations

ESE3 - Imagination 2001 - Continued

Speech Synthesis: Concatenation - II

Must Diphone Synthesis be so Unnatural? ................................. 975
Barry W1, Nielsen C2, Andersen O1
1University of the Saarland, Germany, 2Aalborg University, Denmark

Variable-length acoustic units inference for text-to-speech synthesis .................................................................................. 983
Boeckhoff O
IRISA, Université Rennes 1, ENSSAT, France

Unit Selection for Speech Synthesis Using Splicing Costs with Weighted Finite State Transducers ........................................ 987
Bulyko I, Ostendorf M
University of Washington, USA

Cantonese Text-To-Speech Synthesis Using Sub-Syllable Units .... 991
Law K M, Lee T, Lau W
The Chinese University of Hong Kong, Hong Kong

Session: B43 - Oral

Signal Analysis: Microphone Arrays & Source Localisation

Blind Speech Separation of Moving Speakers Using Hybrid Neural Networks ................................................................. 997
Koutras A, Dermatas E, Kokkinakis G
University of Patras, Greece
For more information: http://www.wcl2.ee.upatras.gr/koutras/on-line.htm

Computationally Efficient Frequency-Domain Combination of Acoustic Echo Cancellation and Robust Adaptive Beamforming1001
Herboldt W, Buchner H, Kellermann W
University Erlangen-Nuremberg, Germany
For more information: http://www.LNT.de/~herbordt/research.html
Non-english Abstract: ne1342.pdf

Calibration of Microphone Arrays for Improved Speech Recognition ................................................................................... 1005
Seltzer M1, Raj B2
1Carnegie Mellon University, USA, 2Compaq Computer Corporation, USA
Improving Simultaneous Speech Recognition in Real
Environments Using Overdetermined Blind Source Separation
Koutras A, Dermatas E, Kokkinakis G
University of Patras, Greece
For more information:

Real-time Sound Source Localization and Separation System and
Its Application to Automatic Speech Recognition
Asano F, Goto M, Iou K, Asoh H
AIST, Japan
For more information:
http://www.etl.go.jp/~asano/research/demo/array2001feb/

Session: B44 - Oral

Speech Recognition and Understanding: Audio-Visual Processing

An Efficient Lipreading Method Using the Symmetry of Lip
Lee J1, Kim F2
1Waseda Univ., Japan, 2Chonnam National University, Korea

Comparing Audio- and A-Posteriori-Probability-Based Stream
Confidence Measures for Audio-Visual Speech Recognition
Heckmann M1, Wild T1, Berthommier F2, Kroschel K2
1Universität Karlsruhe, Germany, 2Institut de la Communication
Parlée, France

Large-vocabulary audio-visual speech recognition by machines
and humans
Potamianos G, Neti C, Iyengar G, Helmut E
IBM T.J. Watson Research Center, USA

Evaluation of an Automatically Obtained Shape and Appearance
Model For Automatic Audio Visual Speech Recognition
Daubias P, Deleglise P
LIUM, France
For more information:
http://www-ic2.univ-lemans.fr/~daubias/recherche.html
Non-english Abstract: ne1844.html

An approach to an Italian Talking Head
Pelachaud C1, Magno-Caldognetto E2, Zmarich C1, Cosi P2
1Università di “La Sapienza”, Italy, 2C.N.R. of Padova, Italy

Session: B45 - Poster

Linguistic Modelling: Language Models - II

Pause Information for Dependency Analysis of Read Japanese
Sentences
Takagi K, Ozeki K
The University of Electro-Communications, Japan

An HMM/N-gram-based Linguistic Processing Approach for
Mandarin Spoken Document Retrieval
Chen B, Wang H-M, Lee L-S
Institute of Information Science, Academia Sinica, Taiwan, ROC

Probabilistic Concept Verification for Language Understanding
in Spoken Dialogue Systems
Lin Y-C, Wang H-M
Industrial Technology Research Institute, Taiwan

Turkish Word Segmentation Using Morphological Analyzer
Kulekci M O1, Ozkan M4
1National Research Institute of Electronics & Cryptology, Turkey,
4Bogazici Univ., Turkey

Thai Grapheme-To-Phoneme Using Probabilistic GLR Parse
Tarsaku P, Sornlertlamvanich V, Thongprasirt R
NECTEC, Thailand

Aligning Prosody and Syntax in Property Grammars
Bluche P, Hirst D
LPL-CNRS, France

From Perceptual Designs to Linguistic Typology and Automatic
Language Identification: Overview and Perspectives
Barkat M, Vasilescu I
University of Lyon 2, France

Session: B46 - Poster

Speech Recognition and Understanding: Noise Robustness - II

Robust Speech Recognition against Packet Loss
Sin M, Chan Y C
Hong Kong University of Science and Technology, Hong Kong

Rapid CODEC Adaptation for Cellular Phone Speech
Recognition
Naito M, Kuroiwa S, Kato T, Shimizu T, Higuchi N
KDDI R&D Laboratories, Japan

A robust front-end for ASR over IP and GSM networks: an
integrated scenario
Gallardo-Antolin A, Pelaez-Moreno C, Diaz-de-Maria F
University Carlos III, Madrid, Spain
Session: C14 - Oral

Speech Recognition and Understanding: Speaker Adaptation

Discriminative Speaker Adaptation with Conditional Maximum Likelihood Linear Regression .............................. 1203
Gunawardana A, Byrne W
The Johns Hopkins University, USA.

What is the Best Type of Prior Distribution for EMAP Speaker Adaptation? ........................................ 1207
Kenny P, Boulianne G, Dumouchel P
Centre de Recherche Informatique de Montreal, Canada

Maximum-Likelihood Affine Cepstral Filtering (MLCAF) Technique for Speaker Normalization ...................... 1211
Kim Y
Stanford University, USA

A Novel Algorithm For Rapid Speaker Adaptation Based On Structural Maximum Likelihood Eigenspace Mapping ... 1215
Zhao B, Hansen J
Univ. of Colorado at Boulder, USA

Evaluation on Unsupervised Speaker Adaptation Based on Sufficient HMM Statistics of Selected Speakers .......... 1219
Yoshizawa S1, Baba A2, Matsunami K3, Mera Y’, Yamada M’, Lee A1, Shikano K1
1Matsushita Electric Industrial Co., Japan, 2Laboratories of Image Information Science and Technology, Japan, 3Nara Institute of Science and Technology, Japan

Session: C15 - Poster

Speech Recognition and Understanding: Adaptation

A Novel Target-Driven MLLR Adaptation Algorithm with Multi-Layer Structure ........................................ 1225
Jia L, Xu B
Institute of Automation, Chinese Academy of Science, P. R. China

Scaled Likelihood Linear Regression for Hidden Markov Model Adaptation .................................. 1229
Wallhoff F, Willert D, Rigoll G
Gerhard-Mercator-University Duisburg, Germany

Fast Adaptation using Constrained Affine Transformations with Hierarchical Priors .......................... 1233
Myrvoll T A1, Paliwal K C2, Svendsen T1
1NTNU, Norway, 2Griffith University, Australia

A Context Adaptation Approach for Building Context Dependent Models in LVCSR .......................... 1237
Liu X, Yuan B, Yan Y
Intel China Research Center, P. R. China

Improving Genericity for Task-Independent Speech Recognition ................................................ 1241
Lefevre F, Gauvain J-L, Lamel L
LIMSI-CNRS, France
Non-english Abstract: ne1817.html

Session: C16 - Poster

Dialogue Systems: Project Descriptions - I

A Posteriori and a Priori Transformations for Speaker Adaptation in Large Vocabulary Speech Recognition Systems 245
Matrouf D, Bellot O, Nocera P, Linares G, Bonastre J-F
LIA, Avignon, France

Bayesian methods for HMM speech recognition with limited training data ........................................ 1249
Purnell D W, Botha E C
University of Pretoria, South Africa

Rapid Speaker Adaptation Using MLLR and Subspace Regression Classes .................................. 1253
Wong K-M, Mak B
The Hong Kong University of Science and Technology, Hong Kong For more information: http://www.cs.ust.hk/~mak

Speaker Adaptation of Output Probabilities and State Duration Distributions for Speech Recognition .......... 1257
Yoma N B, Silva J
University of Chile, Chile
For more information: http://www.die.uchile.cl

Cohorts Based Custom Models for Rapid Speaker and Dialect Adaptation ...................................... 1261
Wu J1, Chang E2
1The University of Hong Kong, Hong Kong, P. R. China, 2Microsoft Research China, P. R. China

Speaker Adaptation of Quantized Parameter HMMs ......... 1265
Vasilache M, Vikki O
Nokia Research Center, Finland

Segmental Eigenvoice for Rapid Speaker Adaptation ...... 1269
Tsao Y, Lee S-M, Lee L-S
National Taiwan University, Taiwan, ROC

Speaker adaptation in an ASR system based on nonlinear dynamical systems .................................. 1273
Warakagoda N D, Johnsen M H
NTNU, Norway

A Multilingual-supporting Dialog System Using a Common Dialog Controller .................................. 1283
Xu Y, Araki M, Niimi Y
Kyoto Institute of Technology, Japan

Graphic platform for designing and developing practical voice interaction systems .......................... 1287
Nouza T, Nouza J
Technical University of Liberec, Czech Republic

Speech Translation for French in the NESPOLE! European Project ........................................ 1291
CLIPS, France
Non-english Abstract: ne1185.pdf
**Lessons from the Development of a Conversational Interface**
Hickey M, St John Brittan P
*HP Laboratories, UK*

**SCANMail: Browsing and Searching Speech Data by Content**
*AT&T Labs - Research, USA, AnswerLogic, USA, Virginia Tech, USA, University of Delaware, USA*

**Multi-Scale Retrieval in MEI: An English-Chinese Translingual Speech Retrieval System**
Lo W-K, Schone P, Meng H
*The Chinese University of Hong Kong, Hong Kong, Department of Defense, USA*
For more information: http://www.cslp.jhu.edu/ws2000/groups/mei/

**Compact Word Graph in Spoken Dialogue System**
Chien S-C, Chang S-C
*Industrial Technology Research Institute, Taiwan*

**Use of Topic Knowledge in Spoken Dialogue Information Retrieval System for Academic Documents**
Kimyama S, Hirose K, Minematsu N
*University of Tokyo, Japan*

**Domain-Independent Spoken Dialogue Platform using Key-Phrase Spotting based on Combined Language Model**
Komatani K, Tanaka K, Kashima H, Kawahara T
*Kyoto University, Japan*

**OASIS Natural Language Call Steering Trial**
*BTexasCT, UK, Bell Labs, USA*

**First steps toward an adaptive spoken dialogue system in medical domain**
Azzini I, Falavigna D, Grettler R, Lanzola G, Orlandi M
*Universita` di Pavia, Italy, ITC-irst, Italy*

**Mokusei: A Telephone-based Japanese Conversational System in the Weather Domain**
*MIT Laboratory for Computer Science, USA / NTT, Japan, NTT, Japan, MIT Laboratory for Computer Science, USA*

**SpeechBuilder: Facilitating Spoken Dialogue System Development**
Glass J, Weinstein E
*MIT Laboratory for Computer Science, USA*
For more information: http://www.sls.lcs.mit.edu/speechbuilder

**Voice-IF: A Mixed-Initiative Spoken Dialogue System for AT&T Conference Services**
*AT&T, USA*

---

**SIGdial - Special Interest Group on Discourse and Dialogue**
Dybkjær L
*Natural Interactive Systems Laboratory, Denmark*

**Integrating Speech Technology in Language Learning: An overview of the activities of InSTIL**
Deleloque P
*University of Aberstwyth Dundee, UK*
For more information: http://www.sls.lcs.mit.edu/speechbuilder

**ISCA SALTMIL SIG: Speech and Language Technology for Minority Languages**
Nadeau C, O’Crónín D, Petek B, Sarasola K, Williams B
*Univ. Politècnica de Catalunya, Spain, Linguistics Institute of Ireland, Ireland, University of Ljubljana, Slovenia, University of the Basque Country, Spain, University of Edinburgh, UK*
For more information: http://www.research.att.com/~walker/eval/eval.html

**The Specificity of French Speech Processing (no proceedings paper)**
Bimbout F, Bonastre J-F
*IRISA, France, LIA, Avignon, France*

---

**Dialogue Systems: Resources**

**Observations on Overlap: Findings and Implications for Automatic Processing of Multi-Party Conversation**
Shriberg E, Stolcke A, Baron D
*SRI International / ICSI, USA, ICSI / U.C. Berkeley, USA*

**Towards SMIL as a Foundation for Multimodal, Multimedia Applications**
Di Fabbrizio G, Nils K, Beckham J
*AT&T Labs Research, USA, University of Wisconsin, USA*

**Anvil - A Generic Annotation Tool for Multimodal Dialogue**
Kipp M
*University of the Saarland, Germany*
For more information: http://www.dfki.de/~kipp/anvil
Non-english Abstract: ne1651.pdf

**DARPA Communicator Dialog Travel Planning Systems: The June 2000 Data Collection**
Walker M
*AT&T Shannon Labs, USA*
For more information: http://www.research.att.com/~walker/eval/eval.html

---

**Speaker Recognition: Features and Transforms**

**Analysis of Speaker Variability**
Huang C, Chen T, Li S, Chang E, Zhou J
*Microsoft Research China, P. R. China, Tsinghua Univ., P. R. China*

**Speaker Recognition by Separating Phonetic Space and Speaker Space**
Nishida M, Ariki Y
*Ryukoku University, Japan*
Improved Data-Driven Generation of Pronunciation Dictionaries
Using an Adapted Word List ......................................... 1433
Wolff M, Eichner M, Hoffmann R
Dresden University of Technology, Germany

Segment-Based Recognition on the PhoneBook Task: Initial
Results and Observations on Duration Modeling ............... 1437
Livescu K, Glass J
MIT Laboratory for Computer Science, USA

Multilingual Text-To-Phoneme Mapping ....................... 1441
Riis S K, Pedersen M W, Jensen K J
Nokia Mobile Phones, Denmark

Pronunciation Variation Analysis with respect to Various
Linguistic Levels and Contextual Conditions for Mandarin
Chinese .............................................................. 1445
Tsai M-Y1, Chou F-C2, Lee L-S3
1National Taiwan University / Applied Speech Technology, Taiwan,
2National Chiao Tung University, Taiwan, 3National Taiwan University,
Taiwan

Hypothesis-driven Accent Discrimination ...................... 1449
Mayfield Tomokiyo L
Carnegie Mellon University, USA

An Approach to Automatic Phonetic Baseform Generation Based
on Bayesian Networks ............................................... 1453
Ma C, Randolph M
Motorola INC, USA

Towards Discriminative Lexicon Optimization ................ 1457
Schramm H, Beyerlein P
Philips Research Laboratories Aachen, Germany

Model Complexity Optimization for Nonnative English Speakers 1461
He X, Zhao Y
University of Missouri, USA

Pronunciation Modeling in Hungarian Number Recognition 1465
Fegyő T, Mihajlik P, Tatai P, Gordos G
Budapest University of Technology and Economics, Hungary

Session: C25 - Poster

Speech Recognition and Understanding:
Pronunciation and Subword Units - II

Dynamic Lexicon Using Phonetic Features .................... 1413
Lee K-T, Welckeens C J
Institut Eurécom, France
Non-english Abstract: ne1092.html

Triphone Tying Techniques combining A-Priori Rules and Data
Driven Methods .................................................... 1417
Ziegenhain U, Bauer J G
Siemens, Germany

Pronunciation modeling and lexical adaptation in mid-size
vocabulary ASR ..................................................... 1421
ten Bosch L, Cremelie N
Lernout & Hauspie Speech Products N.V., Belgium

Estimating Pronunciation Variations from Acoustic Likelihood
Score for HMM Reconstruction ................................. 1425
Liu Y, Fung P
Hong Kong University of Science and Technology, Hong Kong

Breadth-First Search for Finding the Optimal Phonetic
Transcription from Multiple Utterances ...................... 1429
Bisani M, Ney H
Lehrstuhl fuer Informatik VI, RWTH Aachen, Germany

Session: C26 - Poster

Speech Production: Miscellaneous

AMSTIVOC (AMsterdam System for Transcription of Infant
VOCalizations) Applied to Utterances of Deaf and Normally
Hearing Infants .................................................. 1471
Koopmans-van Beinum F J, Clement C J, Van den Dikkenberg-Pot I
University of Amsterdam, the Netherlands

Using Lingualpalatal Contact Patterns to Tune a 3D Tongue
Model ............................................................. 1475
Engwall O
KTH, Sweden
For more information:
http://www.speech.kth.se/multimodal/vocaltract.html

Electromagnetic articulograph (EMA) based on a non-parametric
representation of the magnetic field ............................ 1479
Kaburagi T1, Honda M1
1Kyushu Institute of Design, Japan, 2NTT Communication Science
Laboratories, Japan
Session: D11 - Oral

ESE5 - Existing and Future Corpora: Next Generation Speech Resources

Session Introduction (no proceedings paper) .................. 1524
Draxler C
University of Munich, Germany

Building a Corpus of Natural Speech - and tools for the processing of Expressive Speech .................. 1525
Campbell N
ATR JST/CREST, Japan
For more information: http://www.isd.atr.co.jp/esp

Aspects of Modern Multi-modal/Multi-media Corpora
Exploitation Environments .................................. 1529
Broeder D, Brugman H, Wittenburg P
Max-Planck Institute for Psycholinguistics, the Netherlands

Emerging Requirements for Multi-Modal Annotation and Analysis Tools ........................................... 1533
Bigbee T, Loehr D, Harper L
The MITRE Corporation, USA

Session: D12 - Oral

Dialogue Systems: Project Descriptions - II

SmartKom: Multimodal Communication with a Life-Like Character ............................................. 1547
Walsister W, Reihtinger N, Blocher A
DFKI, Germany
For more information: http://www.smartkom.org

ISIS: A Learning System with Combined Interaction and Delegation Dialogs ...................................... 1551
Meng H1, Chan S F1, Wong Y F1, Chan C C1, Wong Y W1, Fung T Y1, Tsui W C1, Chen K2, Wang L2, Wu T Y2, Li X2, Lee T2, Choi W N1, Ching P C1, Chi H S2
1The Chinese University of Hong Kong, Hong Kong, 2Queen’s University, Canada

Robust Language Understanding in Mipad ......................... 1555
Wang Y-Y
Microsoft Research, USA

The WITAS Multi-Modal Dialogue System I ..................... 1559
Lemon O, Bracy A, Alexander G, Peters S
Stanford University, USA
For more information: http://www-csli.stanford.edu/semlab/witas

Universalizing Speech: Notes from the USI Project ............. 1563
Shriver S, Rosenfeld R, Zhu X, Toth A, Rudnicky A, Flueckiger M
Carnegie Mellon University, USA
For more information: http://www.speech.cs.cmu.edu/usi

Session: D13 - Oral

Signal Analysis: Speech Processing in Car Environments

Use of Real and Contaminated Speech for Training of a Hands-Free In-Car Speech Recognizer .................. 1569
Matassoni M, Omologo M, Svaizer P
ITC-Irst, Italy
For more information: http://poseidon.ist.cnr.it:7117/~ssi/SHINE/
Session: D14 - Oral

Speech Recognition and Understanding: Finite State Transducers for ASR

FST-Based Recognition Techniques for Multi-Lingual and Multi-Domain Spontaneous Speech ........................................... 1591
Hazan T J, Hetherington I L, Park A
MIT Laboratory for Computer Science, USA

A Transducer Approach to Word Graph Generation ............... 1595
Boulia A, Ouellet P, Dumouchel P
Centre de Recherche Informatique de Montréal, Canada

An Efficient Implementation of Phonological Rules using Finite-State Transducers .................................................. 1599
Hetherington I L
MIT Laboratory for Computer Science, USA

A Weight Pushing Algorithm for Large Vocabulary Speech Recognition ..................................................... 1603
Mohri M, Riley M
AT&T Labs - Research, USA

Transducer optimizations for tight-coupled decoding ............ 1607
Sward A
KTH, Sweden

Session: D15 - Poster

Speech Recognition and Understanding: Acoustic Modelling - II

Distinctive Features For Use in an Automatic Speech Recognition System ........................................... 1613
Eude E
IBM, USA

Improved Context-Dependent Acoustic Modeling for Continuous Chinese Speech Recognition ...................... 1617
Zhang J, Zheng F, Li J, Lao C, Zhang G
Tsinghua Univ., P. R. China

Class Definition in Discriminant Feature Analysis ............ 1621
Duchateau J, Demuyynnck K, Van Compernolle D, Wambaq P
Katholieke Universiteit Leuven - ESAT, Belgium
For more information: http://www.esat.kuleuven.ac.be/~spcch

Feature Extraction from Time-Frequency matrices for Robust Speech Recognition ........................................... 1625
Segura J C, Benitez M C, de la Torre A, Rubio A J
1Universidad de Granada, Spain, 2Universidad de Granada, Spain / International Computer Science Institute, USA

Using Spatial Correlation Information in Speech Recognition 1629
Yu P, Wang Z
Tsinghua University, P. R. China

On the Choice of Classes in MCE based discriminative HMM-Training for Speech Recognizers used in the Telephone Environment ........................................................ 1633
Bauer J G
Siemens, Germany

Plosive Spotting with Margin Classifiers .......................... 1637
Keshet J, Chazan D, Bobrovsky B-Z
1Tel Aviv University, Israel, 2IBM - Israel - Science and Technology, Israel

Model Agglomeration for Context-Dependent Acoustic Modeling ................................................................. 1641
Brugnara F
ITC-Irste - Centro per la Ricerca Scientifica e Tecnologica, Italy

Multipass algorithm for acquisition of salient acoustic morphemes ..................................................... 1645
Levit M, Gorin A L, Wright J H
AT&T Laboratories-Research, USA

Rapid Vocal Tract Length Normalization using Maximum Likelihood Estimation ........................................... 1649
Emori T, Shimoda K
NEC Corporation, Japan

Towards the creation of acoustic models for stressed Japanese speech ..................................................... 1653
Okuda K, Matsui T, Nakamura S
ATR Spoken Language Translation Research Laboratories, Japan

Elderly Acoustic Model for Large Vocabulary Continuous Speech Recognition ..................................................... 1657
Baba A', Yoshizawa S', Yamada M', Lee A', Shikano K
1Matsushita Electric Works, Japan, 2Matsushita Electric Industrial Co., Japan, 3Nara Institute of Science and Technology, Japan

A Hybrid Approach to Enhance Task Portability of Acoustic Models in Chinese Speech Recognition ...................... 1661
Zhang J-S, Zhang S-W, Sagisaka Y, Nakamura S
ATR Spoken Language Translation Research Laboratories, Japan

Evaluation of sublexical and lexical models of acoustic disfluencies for spontaneous speech recognition in Spanish ............... 1665
Rodriguez L I, Torres I, Varona A
Universidad del Pais Vasco, Spain

Structural Learning of Dynamic Bayesian Networks in Speech Recognition ..................................................... 1669
Deviren M, Daoudi K
INRIA-LORIA, France

Session: D16 - Poster
### Resources, Assessment and Standards: Assessment Tools & Methodology

**A New Method for Testing Communication Efficiency and User Acceptability of Speech Communication Channels**  ........ 1675  
Van Wijngaarden S, Smelee P, Steeneke H  
TNO Human Factors, the Netherlands  

**Phonetic Transcriptions in the Spoken Dutch Corpus: how to Combine Efficiency and Good Transcription Quality**  .... 1679  
Cucchiarini C, Binnenpoorte D, Goddijn S  
University of Nijmegen, The Netherlands  
For more information: http://lands.let.kun.nl/TSpublic/catia/  

**A Functional Approach to Speech Recognition Evaluation**  1683  
Hutchinson B  
Syrix Speech Systems, Australia  

**Instrumental Derivation of Equipment Impairment Factors for Describing Telephone Speech Codec Degradations**  ........ 1687  
Möller S1, Berger J2  
1Ruhr-Universität Bochum, Germany, 2'T-Nova Deutsche Telekom Innovationsges. mbH Berkom, Germany  

**Julius --- an Open Source Real-Time Large Vocabulary Recognition Engine**  .... 1691  
Lee A1, Kawahara T2, Shikano K5  
1Nara Institute of Science and Technology, Japan, 2Kyoto University, Japan  
For more information: http://www.aist-nara.ac.jp/~ri/  

**Local Refinement of Phonetic Boundaries: A General Framework and Its Application Using Different Transition Models**  .... 1695  
Torre Toledoano D1, Hernández Gómez L A2  
1Telefónica I+D, Spain, 2Univversidad Politécnica de Madrid, Spain  

**Detection of Digital Transmission Systems for Voice Quality Measurements**  ........ 1699  
Ludwig T, Heute U  
Christian-Albrechts University of Kiel, Germany  

**Automatic Segmentation of Recorded Speech into Syllables for Speech Synthesis**  ........ 1703  
Lewis E1, Tatham M2  
1University of Bristol, UK, 2University of Essex, UK  

**Phonetic Events from the Labeling the European Portuguese DataBase for Speech Synthesis, FEUP/IPB-DB**  .... 1707  
Teixeira J P1, Freitas D2, Braga D2, Barros M J2, Latsch V2  
1Instituto Politécnico de Bragança, Portugal, 2IRISA, Université de Rennes 1, ENST, CNRS-LTCI, France  

**Acoustical and topological experiments fo an HMM-based speech segmentation system**  ........ 1711  
Nefi S1, Boeufflard O2  
1France Télécom R&D, DIH/IPS/VMI, France, 2IRISA, Université de Rennes 1, ENSAE, France  

**TelBLASR: An Automatic Speech Recognition Extension for TTS**  1715  
Zhou Q, Zheng J, Lee C-H  
Bell Labs, Lucent Technologies, USA  

---

**Session: D21 - Oral**

**ESE5 - Existing and Future Corpora: Automated Analysis of Speech Resources**

---

**Lower WERs do not guarantee better transcriptions**  .... 1721  
Kessens J M, Strik H  
University of Nijmegen, The Netherlands  
For more information: http://lands.let.kun.nl/staff/kessens.en.html  

**An Elitist Approach to Articulatory-Acoustic Feature Classification**  ........ 1725  
Chang S1, Greenberg S2, Wester M2  
1International Computer Science Institute, USA, 2Nijmegen University, The Netherlands  

**A Dutch Treatment of an Elitist Approach to Articulatory-Acoustic Feature Classification**  .... 1729  
Wester M1, Greenberg S2, Chang S2  
1Nijmegen University, The Netherlands, 2International Computer Science Institute, USA  

---

**Session: D22 - Oral**

**Dialogue Systems: Dialogue Systems and Generation**

**Hybrid Natural Language Generation for Spoken Dialogue Systems**  ........ 1735  
Galley M, Fosler-Lussier E, Potamianos A  
Lucent Technologies, USA  

**The Generation of Speech for a Search Guide**  .... 1739  
Cook N, Benest I  
University of York, UK  

**An Automatic Dialogue System Generator from the Internet Information Contents**  ........ 1743  
Araki M, Ono T, Ueda K, Nishimoto T, Niimi Y  
Kyoto Institute of Technology, Japan  

**Training a Sentence Planner for Spoken Dialog: The Impact of Syntactic and Planning Features**  .... 1747  
Rogati M1, Walker M2, Rambow O3  
1Carnegie Mellon University, USA, 2AT&T Shannon Labs, USA  

---

**Session: D23 - Oral**

**Speaker Recognition: Alternative Trends in Verification - I**

**A Comparative Study of MLP-based Artificial Neural Networks in Text-Independent Speaker Verification Against GMM-based Systems**  .... 1753  
Vivaracho C E1, Ortega-García J2, Alonso L1, Moro Q I1  
1Universidad de Valladolid, Spain, 2Universidad Politécnica de Madrid, Spain, 3Universidad de Salamanca, Spain  

**Enhancing GMM Scores using SVM "Hints"**  .... 1757  
Fine S, Navratil J, Gopinath R  
IBM T.J. Watson Research Center, USA  

**Combining GMM’s with Suport Vector Machines for Text-independent Speaker Verification**  .... 1761  
Kharroubi J, Petrovska-Delacretaz D, Chollet G  
ENST, CNRS-LTCI, France  

**A Text-Independent Speaker Verification System Using Support Vector Machines Classifier**  .... 1765  
Gu Y, Thomas T  
Vocalis Ltd., UK
Session: D24 - Oral

Speech Recognition and Understanding: Speech Understanding

Advances in Automatic Speech Summarization .......... 1771
Hori C, Furui S
Tokyo Institute of Technology, Japan
For more information: http://furui-www.cs.titech.ac.jp/

A Word Graph Interface for a Flexible Concept Based Speech Understanding Framework ................. 1775
Hacihoglu K, Ward W
University of Colorado, USA
For more information: http://communicator.colorado.edu

Comparing grammar-based and robust approaches to speech understanding: a case study ................. 1779
Knight S1, Gorrell G2, Rayner M1, Milward D1, Koeling R1, Lewin I1
1SRI International, UK, 2Netdecisions, UK

Integrating Multiple Knowledge Sources For Improved Speech Understanding .......................... 1783
Aboud S, Scordilis M
University of Miami, USA

Session: D25 - Poster

Speech Recognition and Understanding: Algorithms and Architectures

Classification of Transition sounds with application to Automatic Speech Recognition .................. 1789
Littecher Z1, Chazan D1
1Technion Technology Institute, Israel, 2IBM Israel - Science and Technology, Israel

Gaussian Subtraction (GS) Algorithms for Word Spotting in Continuous Speech ................. 1793
Faizakov A1, Cohen A1, Vaich T1
1Ben-Gurion University, Israel, 2Speech Recognition R&D Center, Israel

Relating Frame Accuracy with Word Error in Hybrid ANN-HMM ASR ........................................ 1797
Shire M L
International Computer Science Institute, USA

A Two-Layer Lexical Tree Based Beam Search in Continuous Chinese Speech Recognition .............. 1801
Zhang G, Zheng F, Wu W
Tsinghua Univ., P. R. China

Automatic Labeling and Digesting for Lecture Speech Utilizing Repeated Speech by Shift CDP .......... 1805
Itoh Y1, Tanaka K2
1Iwate Prefectural University, Japan, 2Electrotechnical Laboratory, Japan

Improved Phoneme-History-Dependent Search for Large-Vocabulary Continuous-Speech Recognition . 1809
Hori T1, Noda Y1, Matsunaga S1
1NIT Cyber Space Laboratories, Japan, 2NIT Communications Corporation, Japan

Comparison of MFCC and PLP Parameterizations in the Speaker Independent Continuous Speech Recognition Task 1813
Psutka J, Muller L, Psutka J V
University of West Bohemia, Czech Republic

N-best List Generation using Word and Phoneme Recognition Fusion ..................................... 1817
Pusateri E1, Van Thong J-M1
1MIT Spoken Language Group, USA, 2Compaq Cambridge Research Laboratory, USA

A One Pass Semi-dynamic Network Decoder based on Language Model Network .......................... 1821
Ahn D-H, Chung M
Sogang University, Korea

Improving Automatic Speech Recognition Using Tangent Distance ........................................... 1825
Macherey W, Keyers D, Dahmen J, Nye H
RWTH Aachen, University of Technology, Germany

N-best Speech Hypotheses Reordering Using Linear Regression ................................................ 1829
Chotimongkol A, Rudnicky A
Carnegie Mellon University, USA

Low-Resource Hidden Markov Model Speech Recognition ....................................................... 1833
IBM, USA

Speech recognition at multiple sampling rates .......... 1837
Hirsch H-G1, Hellwig K1, Dobler S1
1University of Applied Sciences Niederrhein, Germany, 2Ericsson Eurolab, Spain

Support Vector Machine with Dynamic Time-Alignment Kernel for Speech Recognition ............... 1841
Shimodaira H1, Noma K-I1, Nakai M1, Sagayama S1
1Japan Advanced Institute of Science and Technology, Japan, 2University of Tokyo, Japan

Efficient Scalable Speech Compression for Scalable Speech Recognition .................................... 1845
Srinivasamurthy N, Ortega A, Narayanan S
University of Southern California, USA

Session: D26 - Poster

Signal Analysis: Speech Enhancement and Noise Processing

Voice Activity Detection in Noisy Environments .................. 1851
Studeramn F1, Stahl V2, Rose G2
1University of Duisburg, Germany, 2Philips Research Laboratories, Germany

An Improved Wavelet-Based Speech Enhancement System ............................................. 1855
Sheikhzadeh H1, Abatalbi H R2
1Amirkabir University of Technology, Iran / DspFactory Ltd., Canada, 2Amirkabir University of Technology, Iran

Enhancing Distributed Speech Recognition with Back-End Speech Reconstruction ................. 1859
Ramabadran T, Meunier J, Jasiuk M, Kushner B
Motorola, USA
Implementation Effective One-Channel Noise Reduction System
Tihelka J, Sovka P
FEE CTU in Prague, Czech Republic

Efficient Speech Enhancement by Diffusive Gain Factors (DGF)
Kim H-G, Obermayer K, Bode M, Ruwisch D
"Technical University of Berlin, Germany, Cortologic, Germany"

Correction of the Voice Timbre Distortions on Telephone Network
Mahé G, Gilloire A
France Télécom R&D / DIH / IPS, France

Speech Enhancement based on IMM with NPHMM
Lee Y, Lee J, Lee K Y, Shirai K
"Soongsil University, Korea, Waseda Univ., Japan"

Speech Recognition under Musical Environments Using Kalman Filter and Iterative MLLR Adaptation
Fujimoto M, Ariki Y
Ryukoku University, Japan

Dual Channel Speech Enhancement using Coherence Function and MDL-based Subspace Approach in Bark Domain
Vetter R, Renevey P, Krauss J
CSEM, Switzerland

Entropy Based Voice Activity Detection in Very Noisy Conditions
Renevey P, Drygajlo A
"CSEM, Switzerland, EPFL, Switzerland"

Discrimination between speech and music based on a low frequency modulation feature
Karnebäck S
KTH, Sweden

Credibility Proof for Speech Content and Speaker Verification by Fragile Watermarking with Consecutive Frame-Based Processing
Cheng Y-W, Lee L-S
National Taiwan University, Taiwan

Map Estimation for On-line Noise Compensation of Time Trajectories of Spectral Coefficients
Potamitis I, Fakotakis N, Kokkinakis G
University of Patras, Greece

A New Method for Speech Denoising and Robust Speech Recognition Using Probabilistic Models for Clean Speech and for Noise
Attias H, Deng L, Acero A, Platt JC
Microsoft Research, USA

Designing very compact decision trees for grapheme-to-phoneme transcription
Kienappel A, Kneser R
Philips Research Laboratories, Germany

Using Machine Learning Techniques for Grapheme to Phoneme Transcription
Mana F, Massimino P, Pacchiotti A
Loquendo, Vocal Technology and Services, Italy

Knowledge of Language Origin Improves Pronunciation Accuracy of Proper Names
Liitjos A F, Black A W
Carnegie Mellon Univ., USA

On the pronunciation of acronyms in French and in Italian
Boula de Mareuil P, Floricic F
"Elan TTS, France, ERSS, France"

Session: D33 - Oral
Signal Analysis: Speech Enhancement
Enhancement of Noisy Speech by Using Improved Global Soft Decision
Samsung AIT, Korea

Enhancement of Speech Using Bark-Scaled Wavelet Packet Decomposition
Cohen I
Lamar Signal Processing Ltd., Israel

A New Approach for Wavelet Speech Enhancement
Bahoura M, Rouat J
Université du Québec à Chicoutimi, Canada
For more information: http://wwwdss.uqac.uquebec.ca/ermetis

Speech/Noise-Dominant Decision for Speech Enhancement
Yoon S, Yoo C D
Korea Advanced Institute of Science and Technology (KAIST), Korea
For more information: http://fft1.kaist.ac.kr/~neo/speech.html

Session: D34 - Oral
Speech Recognition and Understanding: Discriminative Training
An MCE based Classification Tree Using Hierarchical Feature-Weighting in Speech Recognition
Wang F, Zheng F, Wu W
Tsinghua University, P. R. China

Selective MCE Training Strategy in Mandarin Speech Recognition
Zhou J, Chang E, Chao H
Microsoft Research China, P. R. China

Discriminative Disfluency Modeling for Spontaneous Speech Recognition
Wu C-H, Yan G-L
National Cheng Kung University, Taiwan, ROC
Comparative Analysis for Data-Driven Temporal Filters Obtained Via Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) In Speech Recognition 1959
Hung J-W1, Wang H-M1, Lee L-S2
1Institute of Information Science, Academia Sinica, Taiwan, 2National Taiwan University, Taiwan

Session: D35 - Poster
Speech Coding: Advances in Speech Coding

Coding method for successive pitch periods 1965
Heikkonen A, Ruoppola V T, Pietila S
Nokia Research Center, Finland

Objective Evaluation of Methods for Quantization of Variable-Dimension Spectral Vectors in WI Speech Coding 1969
Nurminen J1, Heikkonen A2, Saarinen J3
1Tampere University of Technology, Finland, 2Nokia Research Center, Finland

Squared Error as a Measure of Phase Distortion 1973
Pohloth H, Kleijn W B
KTH, Sweden
For more information: http://www.speech.kth.se/speech_processing/

Non-Linear Predictive Vector Quantization of speech 1977
Faundez-Zanuy M
Escola Universitaria Politecnica de Mataro, Spain
For more information: http://www.eupmt.es/veu

A Variable Rate Hybrid Coder Based on a Synchronized Harmonic Excitation 1981
Katugampala N, Kondoza A
University of Surrey, UK

A Hybrid Sub-Band Sinusoidal Coding Scheme 1985
Ho M-S, Molyneux D J, Cheetham B M G
University of Manchester, UK

Low rate speech coding incorporating Simultaneously Masked Spectrally Weighted Linear Prediction 1989
Luaski J, Burnett I S, Ritzi C H
University of Wollongong, Australia

Narrowband Perceptual Audio Coding: Enhancements for Speech 1993
Najaf-Zadeh H, Kabal P
McGill University, Canada
For more information: http://www.tsp.ece.mcgill.ca/

Techniques for high-quality ACELP coding of wideband speech 2007
Bessette B1, Lefebvre R2, Salami R3, Jelinek M1, Vainio J2, Rotola-Pukkila J1, Mikkola H1, Jarvinen K3
1University of Sherbrooke, Canada, 2VoicAge Corporation, Canada, 3Nokia Research Center, Finland

Wideband ACELP at 16 kb/s with Multi-band Excitation 2001
Pujalte S, Moreno A
Universitat Politècnica de Catalunya, Spain

Wideband Speech Coding Algorithm with Application of Discrete Wavelet Transform to Upper Band 2005
Lee S W, Bae K S
Kyungpook National University, Korea

A Switched DPCM/Subband Coder for Pre-echo Reduction 2009
Satheesh S, Sreenivas T V
Indian Institute of Science, India

A Generalized Multistage VQ Approach for Spectral Magnitude Quantization 2013
Ettemoglu C O, Cuperman V
University of California, Santa Barbara, USA

Efficient Implementation of ITU-T G.723.1 Speech Coder for Multichannel Voice Transmission and Storage 2017
Jung S K, Park Y C, Yoon S W, Kim K T, Youn D H
Yonsei University, Korea

Session: D36 - Poster
Resources, Assessment and Standards: Corpora

"CU-Move" : Analysis & Corpus Development for Interactive In-Vehicle Speech Systems 2023
Univ. of Colorado - Boulder, USA

Multimedia Data Collection of In-Car Speech Communication 2027
Kawaguchi N, Matsubara S, Takeda K, Itakura F
Nagoya University, Japan

The U.S. SpeechDat-Car Data Collection 2031
Heeman P A, Cole D, Cronk A
Oregon Graduate Institute, USA

Word Unit Based Multilingual Comparative Analysis of Text Corpora 2035
Németh G, Zankó C
Budapest University of Technology and Economics, Hungary
For more information: http://speechlab.ttt.bme.hu

Creating a European English Broadcast News Transcription Corpus and System 2039
Backfried G, Hecht R, Loots S, Pfannnerer N, Riedler J, Schiefer C
Sail Labs, Austria

The Nespoli! VoIP Dialogue Database 2043
Burger S1, Besacier L2, Coletti P3, Metze F1, Morel C1
1Carnegie Mellon University, USA, 2Université Joseph Fourier, France, 3Istituto Trentino di Cultura - Centro per la Ricerca Scientifica e Tecnologica, Italy, 4University of Karlsruhe, Germany

Design of Speech Corpus for Text-to-Speech Synthesis 2047
Matousek J, Psotka J, Kruta J
University of West Bohemia in Pilsen, Czech Republic

The IFA Corpus: a Phonemically Segmented Dutch "Open Source" Speech Database 2051
Van Son R J J H1, Binnenpoorte D2, Van den Heuvel H2, Pols L C W3
1University of Amsterdam, the Netherlands, 2Nijmegen University, the Netherlands
For more information: http://www.fon.hum.uva.nl/IFAcorpus/

African Speech Technology (AST) Telephone Speech Databases: Corpus Design and Contents 2055
Louw P H1, Rous J C2, Boep Boe2
1University of Stellenbosch, South Africa, 2University of Pretoria, South Africa
For more information: http://www.ast.sun.ac.za
SpeechDat-E: Five Eastern European Speech Databases for Voice-Operated Teleservices Completed ........................................ 2059
Van den Heuvel H1, Boudy J1, Bakcsi Z1, Cernocky J1, Galunov V1, Kochanina J1, Majewski W1, Pollak P1, Rusko M1, Sadowski J1, Staronezicz P1, Tropf H1
1University of Nijmegen, the Netherlands, 2Lernout & Hauspie France, France, 3Budapest University of Technology and Economics, Hungary, 4Brno University of Technology, Czech Republic, 5AudiTech, Ltd, St.Petersburg, Russia, 6ITA, Wroclaw University of Technology, Poland, 7Czech Technical University in Prague, Czech Republic, 8Institute of Control Theory and Robotics, Slovak Academy of Sciences, Slovakia, 9Wroclaw University of Technology, Poland, 10Siemens, Germany

Concordancing for Parallel Spoken Language Corpora . . . 2063
Gibbon D1, Trippel T1, Sharoff S2
1Universität Bielefeld, Germany, 2Russian Research Institut for Artificial Intelligence / Universität Bielefeld, Germany

Large Broadcast News and Read Speech Corpora of Spoken Czech ............................................ 2067
Psatka J1, Radova V1, Muller L1, Matousek J1, Ircing P1, Graff D1
1University of West Bohemia in Pilsen, Czech Republic, 2University of Pennsylvania, USA

Development of Russian Lexical Databases, Corpora and Supporting Tools for Speech Products ............................. 2071
Yablonsky S
St. Petersburg State Transport University, Russia

Constructing a segment database for Greek time domain speech synthesis ......................................................... 2075
Fotinea S-E, Tambouratzis G, Carayannis G
Institute for Language and Speech Processing, Greece

Session: D41 - Demonstrations
ESE6 - Education Arena - Continued ............................................. 2080

Session: D42 - Oral
Resources, Assessment and Standards: Assessment Methodology

Subjective Assessment of Speech-System Interface Usability 2083
Hone K1, Graham R1
1Brunel University, UK, 2Motorola, UK
For more information: http://www.brunel.ac.uk/~cssksh/

An Objective Measure for Estimating MOS of Synthesized Speech ................................................................. 2087
Cha M, Peng H
Microsoft Research China, P. R. China

Comparing the performance of two CSRs: How to determine the significance level of the differences ...................... 2091
Strik H, Cucchiariini C, Kessens J
Univ. of Nijmegen, the Netherlands
For more information: http://lands.let.kun.nl/~TSpública/str/ict/

Prediction of Low Recognition Rate Words for Isolated Word Recognition System ........................................ 2095
Tetashima R, Hoshino H, Wakita T
Toyota Central R&D Labs., Japan

An Objective Measure for Assessment of the Concatenative TTS Segment Inventories ......................................... 2099
Batusek R
Masaryk University, Czech Republic

Session: D43 - Oral
Speech Recognition and Understanding: Confidence Measures

Word Level Confidence Annotation using Combinations of Features ................................................................. 2105
Zhang R, Rudnicky A
Carnegie Mellon University, USA

A Boosting Approach for Confidence Scoring .................... 2109
Moreno P J, Logan B, Raj B
Compaq Cambridge Research Lab., USA

On Combining Confidence Measures for Improved Rejection of Incorrect Data .................................................. 2113
Charlet D, Mercier G, Jouvet D
France Télécom R&D, France

Improved Word Confidence Estimation using Long Range Features ................................................................. 2117
Palmer D1, Ostendorf M2
1The MITRE Corporation, USA, 2University of Washington, USA

Is This Conversation on Track? ................................. 2121
Carpenter P, Jin C, Wilson D, Zhang R, Bohus D, Rudnicky A
Carnegie Mellon University, USA

Session: D44 - Oral
Speech Recognition and Understanding: Language Modelling

Automatic N-gram Language Model Creation from Web Resources ................................................................. 2127
Nisimura R1, Komatsu K2, Kuroda Y1, Nagatomo K1, Lee A1, Saruwatari H1, Shikano K1
1Nara Institute of Science and Technology, Japan, 2Laboratories of Image Information Science and Technology, Japan, 3TIS Inc., Japan

On Integrating the Lexicon with the Language Model ........ 2131
Caseiro D A, Trancoso I
INE’S-ID/IST, Portugal
Non-english Abstract: ne1617.pdf

Back-off smoothing evaluation over syntactic language model ................................................................. 2135
Varona A, Torres I
Facultad de Ciencias. UPV-EHU., Spain

An Online Incremental Language Model Adaptation Method ................................................................. 2139
Wu G, Zheng F, Jin L, Wu W
Tsinghua University, P. R. China
Using Boosting and POS Word Graph Tagging to Improve Speech Recognition .......................... 2143
Samuelsson C, Hieronymus P
1Xerox Research Centre Europe, France, 2Research Institute for Advanced Computer Science, USA

Dialogue Systems: Techniques and Strategies

Robust Parsing in Spoken Dialogue Systems .......... 2149
Yan P, Zheng F, Xu M
Tsinghua University, P. R. China
Non-english Abstract: ne24.pdf

A Theme Structure Method for the Ellipsis Resolution ...... 2153
Huang Y, Zheng F, Su Y, Li F, Wu W
Tsinghua University, P. R. China

Deriving Document Structure from Prosodic Cues ........ 2157
Haase M1, Knechtau W2, Mohler G3, Stenzel G4
1Universität Stuttgart, Germany, 2IBM Deutschland Entwicklung, Germany
For more information: http://www.ims.uni-stuttgart.de/~knechtau

Methodology for Dialogue Design in Telephone-Based Spoken Dialogue Systems: a Spannish Train Information System .. 2165
UPM., Spain
For more information: http://www-gth.die.upm.es
Non-english Abstract: ne1080.pdf

Spoken Dialogue Management as Planning and Acting under Uncertainty .................................. 2169
Zhang B1, Cai Q1, Mao J1, Chang E1, Guo B2
1Univ. of Sci. & Tech. of China, 2Tsinghua University, P. R. China, 3Microsoft Research China, P. R. China

Modeling of Conversational Strategy for the Robot Participating in the Group Conversation .......... 2173
Matussaka Y, Fuji S, Kobayashi T
Waseda Univ., Japan
For more information: http://www.tlk.elec.waseda.ac.jp/

Supporting the Construction of a User Model in Speech-only Interfaces by Adding Multi-modality ........ 2177
Terken J, Riele, te S
Technische Universiteit Eindhoven, the Netherlands
For more information: http://www.ipo.tue.nl/projects/matis/

A Word- and Turn-Oriented Approach to Exploring the Structure of Mandarin Dialogues .................. 2181
Tseng S-C
Institute of Linguistics, Academia Sinica, Taiwan

A Rule Based Approach to Extraction of Topics and Dialog Acts in a Spoken Dialog System ............. 2185
Niiyi Y, Oka T, Nishimoto T, Araki M
Kyoto Institute of Technology, Japan

Agent-based Error Handling in Spoken Dialogue Systems .......... 2189
Turunen M, Hakulinen J
Univ. of Tampere, Finland
For more information: http://www.cs.uta.fi/research/hci/SUI/

Iterative Implementation of Dialogue System Modules .......... 2193
Degerstedt L, Jonsson A
Linköping University, Sweden

Oppermann D, Schiel F, Steininger S, Beringer N
University of Munich, Germany
For more information: http://smartkom.dfki.de

Automatic Analysis of Real Dialogues and Generating of Training Corpora .................................. 2201
Schwarz J1, Matousek V2
1Technical University of Dresden, Germany, 2University of West Bohemia in Plsen, Czech Republic

Natural Language Understanding Using Statistical Machine Translation ......................................... 2205
Macherey K, Och F J, Ney H
RWTH Aachen, University of Technology, Germany

Improvements in Audio Processing and Language modeling in the CU Communicator ...................... 2209
University of Colorado at Boulder, USA
For more information: http://csdr.colorado.edu
Non-english Abstract: ne1636.pdf

Dialogue Session Management Using VoiceXML .......... 2213
Tsai A, Pargellis A N, Lee C-H, Olive J P
Bell Labs, Lucent Technologies, USA

Ambiguity Representation and Resolution in Spoken Dialogue Systems ........................................ 2217
Ammicht E, Potamianos A, Fosler-Lussier E
Bell Labs, Lucent Technologies, USA

Speech Synthesis: Miscellaneous

Feature Extraction by Auditory Modeling for Unit Selection in Concatenative Speech Synthesis .................. 2223
Tsuzaki M
ATR Spoken Language Translation Research Laboratories, Japan

Perceptual Cost Functions for Unit Searching in Large Corpus-based Text-to-Speech ......................... 2227
Lee M
Bell Labs, Lucent Technologies, USA

Pruning of Redundant Synthesis Instances Based on Weighted Vector Quantization ......................... 2231
Kim S1, Lee Y1, Hirose K2
1Electronics and Telecommunications Research Institute, Korea, 2University of Tokyo, Japan

Using Real Words for Recording Diphones ................. 2235
Fitt S
University of Edinburgh, UK
For more information: http://www.cstr.ed.ac.uk/projects/unisyn.html
Application of the Trended Hidden Markov Model to Speech Synthesis ................................................ 2239
Dines J D S, Sridharan S, Moody M
RCSAVT, QUT, Australia

Two Features to Check Phonetic Transcriptions in Text to Speech Systems ........................................ 2243
Sandri S, Zovato E
Loquendo, Vocal Technology and Services, Italy

Text-to-Speech Scripting Interface for Appropriate Vocalisation of e-Texts ........................................ 2247
Xydas G, Kouroupetroglou G
University of Athens, Greece

Representation of Large Lexica Using Finite-State Transducers for the Multilingual Text-to-Speech Synthesis Systems . . . 2251
Rojc M, Kacic Z
University of Maribor, Slovenia

Corpus-Based Synthesis of Fundamental Frequency Contours Based on A Generation Process Model .................. 2255
Hirose K1, Eto M1, Mienmutu N1, Sakurai A1
1University of Tokyo, Japan, 2Tsukuba R&D Center, Texas Instruments Japan

Corpus-Based Database of Residual Excitations Used for Speech Reconstruction from MFCCs ...................... 2259
Tychtl Z, Ptakka J
University of West Bohemia, Czech Republic

Mixed Excitation for HMM-based Speech Synthesis . . . . 2263
Yoshimura T1, Tokada K1, Masuko T1, Kobayashi T1, Kitamura T1
1Nagoya Institute of Technology, Japan, 2Tokyo Institute of Technology, Japan
For more information: http://kt-lab.ics.nitech.ac.jp/~yossie/TTS

Aperiodicity Control in ARX-Based Speech Analysis-Synthesis Method .................................................. 2267
Ohtsuka T, Kasaya H
Utsunomiya University, Japan
For more information: http://www.klab.ee.utsunomiya-u.ac.jp

Generalized Source-Filter Structures for Speech Synthesis 2271
Karjalainen M, Paatero T
Helsinki University of Technology, Finland
For more information: http://www.acoustics.hut.fi/~mak/

The Speech Synthesis Environment and Parametric Modeling of Coarticulation ...................................... 2275
Mikolaj W
 Adam Mickiewicz University, Poland

Session: E11 - Oral

ESE7 - Integration of Phonetic Knowledge in Speech Technology: Experiments and Experiences . . . . . . 2281
Carson-Berndsen J, Walsh M
University College Dublin, Ireland

Prosodic models, automatic speech understanding, and speech synthesis: towards the common ground. . . . 2285
Batliner A1, Mobius B2, Mohler G2, Schweitzer A2, Noth E1
1University of Erlangen-Nuremberg, Germany, 2University of Stuttgart, Germany

Introducing Phonetically Motivated Information into CSR 2289
Christensen H, Lindberg B, Andersen O
Aalborg University, Denmark

Integrating Contextual Phonological Rules in a Large Vocabulary Decoder ................................................. 2293
Gravier G1, Yvon F2, Jacob B1, Bimbot F2
1IRISA, France, 2ENST Paris, France, 3LIUM, France
Non-english Abstract: ne1820.pdf

Automatic Learning of Finite State Automata for Pronunciation Modeling .............................................. 2297
Pastor-i-Gadea M, Casacuberta F
Universitat Politècnica de València, Spain

Session: E12 - Oral

Speech Coding: Wideband Speech Coding

AMR Wideband codec - leap in mobile communication voice quality .................................................... 2303
Rotola-Pikkala J1, Vainio J2, Mikko H1, Jarvinen K1, Bessette B2
1Nokia Research Center, Finland, 2University of Sherbrooke, Canada, 3VoiceAge Corporation, Canada

Combined Speech and Audio Coding with Bit Rate and Bandwidth Scalability ...................................... 2307
Furrugia M, Kondoz A M
University of Surrey, UK

Joint Speech and Audio Coding Combining Sinusoidal Modeling and Wavelet Packets ................................ 2311
Fek M1, Varkonyi-Koczys A2, Boucher J-M2
1Budapest University of Technology and Economics, Hungary, 2ENST de Bretagne, France
For more information: http://www-sc.enst-bretagne.fr/~fek/eurospeech01

Temporal Decomposition: A Promising Approach to Low Rate Wideband Speech Compression ............ 2315
Ritz C H, Burnett I S
University of Wollongong, Australia

Wideband LSF Quantization by Generalized Voronoi Codes 2319
Ragot S, Lahdili H, Lefebvre R
Univ. Sherbrooke, Canada
For more information: http://www-edu.gel.usherb.ca/ragot

Session: E13 - Oral

Dialogue Systems: Techniques and Strategies

Learning of User Formulations for Business Listings in Automatic Directory Assistance ......................... 2325
Popovici C1, Andorno M1, Laface P1, Fissore L1, Negra M2, Vair C1
1Politecnico di Torino, Italy, 2Loquendo, Italy

Mathematical modeling of Spoken Human Machine dialogues including erroneou s confirmations .......... 2329
Louloudis D1, Tsapanoglou A2, Fakotakis N1, Kokkinakis G1
1University of Patras, Greece, 2Knowledge S.A., Greece

Limited Enquiry Negotiation Dialogues ........................................ 2333
Levin I
Netdecisions Ltd, UK
A Comparison of some Different Techniques for Vector Based Call-Routing 2337  
Cox S1, Shahshahani B2  
1University of East Anglia, UK, 2Nuance Communications, USA

Architecture for adaptive multimodal dialog systems based on VoiceXML 2341  
Nikfeld G, Finan R, Pucher M  
1Telecommunications Research Center Vienna, Austria, 2Mobilkom Austria AG, Austria  
For more information: http://speech.ftw.at/~nikfeld

Session: E14 - Oral

Speech Recognition and Understanding: Robust ASR

Separating speaker and environment variabilities for improved recognition in non-stationary conditions 2347  
Rigazio L, Nguyen P, Kryze D, Junqua J-C  
PSTL, USA

Robust speech recognition techniques applied to a speech in noise task 2351  
Rose R C1, Kim H K1, Hindle D2  
1AT&T Labs - Research, USA, 2AnswerLogic Inc., USA

Minimax Classification With Parametric neighborhoods For Noisy Speech Recognition 2355  
Affy M, Shahshani O, Lee C-H  
Bell Laboratories, Lucent Technologies, USA

Maximum Likelihood Non-linear Transformation for Environment Adaptation in Speech Recognition 2359  
Padmanabham M, Dharanipragada S  
IBM T. J. Watson Research Center, USA

A Study of Speech Coding Parameters in Speech Recognition 2363  
Turunen J J1, Vlaj D2  
1Tampere University of Technology, Finland, 2University of Maribor, Slovenia

Session: E15 - Poster

Applications: Miscellaneous Applications

Some Practical Considerations in the Deployment of a Wireless-Communication Interactive Voice Response System 2369  
Garcia-Mateo C, Docio-Fernandez L, Cardenal-Lopez A  
University of Vigo, Spain

Caller Identification for the SCANMail Voicemail Browser 2373  
Rosenberg A1, Hirschberg J1, Bacchiani M1, Parthasarathy S1, Isenhour P1, Stead L1  
1AT&T Labs-Research, USA, 2Virginia Tech, USA

Extractive Summarization of Voicemail using Lexical and Prosodic Feature Subset Selection 2377  
Koumpis K, Renals S, Niranjan M  
University of Sheffield, UK  
For more information: http://www.dcs.shef.ac.uk/spandh/projects/speechsum.html

Business Listings in Automatic Directory Assistance 2381  
Scharenborg G, Sturm J, Boves L  
University of Nijmegen, the Netherlands

 EuTrans: a Speech-to-Speech Translator Prototype 2385  
Pastor-i-Gadea M, Sanchis A, Casacuberta F, Vidal E  
Institut Tecnologic d'Informatica Universitat Politencica de Valencia, Spain

Speech Recognition over NetMeeting Connections 2389  
Metze F, McDonough J, Soltan H  
University of Karlsruhe, Germany

DIARCA: A Component Approach to Voice Recognition 2393  
Díaz Martín J C1, García Zapata J L1, Rodríguez García J M1, Álvarez Salgado J F1, Espada Bueno P1, Gómez Vilda P2  
1Universidad de Extremadura, Spain, 2Universidad de Politencica de Madrid, Spain

The mVprotek : m-commerce Voice verification system 2397  
Kyung Y1, Jung J, Sohn S M1, Chun H J1, Moon S Y2, Kim M H1, Sull W H1  
1SK Telecom Platform R&D Center, Korea, 2Infinity Telecom ICN Lab, Korea

Real-time Multilingual Communication by means of Prestored Conversational Units 2401  
Alm N1, Iwabuchi MF, Andresen P N1, Nakamura K2, Murray IR3  
1University of Dundee, UK, 2Hiroshima University, Japan, 3Kagawa University, Japan

Writing script-based dialogues for AAC 2405  
University of Dundee, UK  
For more information: http://www.computing.dundee.ac.uk/research/rs_funded_projects_view.asp?id=66

Communication Aid for non-vocal people using corpus-based concatenative speech synthesis 2409  
Iida A1, Sakurada Y1, Campbell N1, Yasumura M1  
1Keio University, Japan, 2ATR International, Japan

Social Effects on Vocal Rate with Echic Mimicry Using Prosody-only Voice 2413  
Suzuki N1, Kakehi K1, Takeuchi Y1, Okada M1  
1ATR MI&C / Nagoya University, Japan, 2Nagoya University, Japan, 3ATR MI&C, Japan

Everyday Life Sounds and Speech Analysis for a Medical Telemonitoring System 2417  
Castelli E, Istrate D  
CLIPS/IMAG, France

Speaking While Driving - Preliminary Results on Spellings in the German SpeechDat-Car Database 2421  
Druker C1, Bengler K2  
1University of Munich, Germany, 2BMW Group, Munich, Germany

Session: E16 - Poster

Signal Analysis: Pitch and Speech Analysis

Efficient Periodicity Extraction Based on Sine-Wave Representation and its Application to Pitch Determination of Speech Signals 2427  
Chazan D1, Tzur M1, Hoory R1, Cohen G1  
1IBM Research, Israel, 2BIT Innovation Technologies, Israel
**Session: E21 - Oral**

**ESE7 - Integration of Phonetic Knowledge in Speech Technology: Is Phonetic Knowledge any use? Panel discussion**

From Here to Utility - Melding Phonetic Insight with Speech Technology

Greenberg S
International Computer Science Institute, USA
For more information: http://www.icsi.berkeley.edu/~steveng

**Session: E22 - Oral**

**Speech Coding: Speech Transmission Systems**

Speech Quality Measure for VoIP using Wavelet based Bark

Park S-W, Park Y-C, Youn D-H
Yonsei University, Korea

A Proposed Method for Measuring Language Dependency of Narrow Band Voice Coders

Van Wijngaarden S, Steeneken H
TNO Human Factors, the Netherlands

An Efficient Transcoding Algorithm For G.723.1 And G.729A

Yoon S W, Jung S K, Park Y C, Youn D H
Yonsei University, Korea

**Session: E23 - Oral**

**Speaker Recognition: Alternative Trends in Verification - II**

A Segmental Mixture Model for Speaker Recognition

Stapert R P, Mason J S
University of Wales Swansea, UK
For more information: http://galilee.swan.ac.uk

Tree Based Score Computation for Speaker Verification

Blouet R, Bimbot F
IRISA (CNRS & INRIA), France
Non-english Abstract: ne1845.pdf

Phonetic Speaker Recognition

Perez-Cordoba J L, Rubio A J, Peinado A M, de la Torre A
Universidad de Granada, Spain
Non-english Abstract: ne1553.pdf

**Session: E24 - Oral**

**Speech Recognition and Understanding: Rhythm and Timing in ASR**

A Segmental Mixture Model for Speaker Recognition

Stapert R P, Mason J S
University of Wales Swansea, UK
For more information: http://galilee.swan.ac.uk

Tree Based Score Computation for Speaker Verification

Blouet R, Bimbot F
IRISA (CNRS & INRIA), France
Non-english Abstract: ne1845.pdf

Phonetic Speaker Recognition

Perez-Cordoba J L, Rubio A J, Peinado A M, de la Torre A
Universidad de Granada, Spain
Non-english Abstract: ne1553.pdf

**ESE7 - Integration of Phonetic Knowledge in Speech Technology: Is Phonetic Knowledge any use? Panel discussion**

From Here to Utility - Melding Phonetic Insight with Speech Technology

Greenberg S
International Computer Science Institute, USA
For more information: http://www.icsi.berkeley.edu/~steveng

**Session: E22 - Oral**

**Speech Coding: Speech Transmission Systems**

Speech Quality Measure for VoIP using Wavelet based Bark

Park S-W, Park Y-C, Youn D-H
Yonsei University, Korea

A Proposed Method for Measuring Language Dependency of Narrow Band Voice Coders

Van Wijngaarden S, Steeneken H
TNO Human Factors, the Netherlands

An Efficient Transcoding Algorithm For G.723.1 And G.729A

Yoon S W, Jung S K, Park Y C, Youn D H
Yonsei University, Korea

**Session: E23 - Oral**

**Speaker Recognition: Alternative Trends in Verification - II**

A Segmental Mixture Model for Speaker Recognition

Stapert R P, Mason J S
University of Wales Swansea, UK
For more information: http://galilee.swan.ac.uk

Tree Based Score Computation for Speaker Verification

Blouet R, Bimbot F
IRISA (CNRS & INRIA), France
Non-english Abstract: ne1845.pdf

Phonetic Speaker Recognition

Perez-Cordoba J L, Rubio A J, Peinado A M, de la Torre A
Universidad de Granada, Spain
Non-english Abstract: ne1553.pdf

**Session: E24 - Oral**

**Speech Recognition and Understanding: Rhythm and Timing in ASR**

A Segmental Mixture Model for Speaker Recognition

Stapert R P, Mason J S
University of Wales Swansea, UK
For more information: http://galilee.swan.ac.uk

Tree Based Score Computation for Speaker Verification

Blouet R, Bimbot F
IRISA (CNRS & INRIA), France
Non-english Abstract: ne1845.pdf

Phonetic Speaker Recognition

Perez-Cordoba J L, Rubio A J, Peinado A M, de la Torre A
Universidad de Granada, Spain
Non-english Abstract: ne1553.pdf
An investigation of modelling aspects for rate-dependent speech recognition .............................................. 2527
Wrede B, Fink G A, Sagerer G Bielefeld University, Germany

Speaking Rate Dependent Acoustic Modeling for Spontaneous Lecture Speech Recognition ........................ ... 2531
Nanjio H, Kato K, Kawahara T Kyoto University, Japan

Analysis of N-Best Output Hypotheses for Fast Speech in Large Vocabulary Continuous Speech Recognition ........ 2535
Fábián T, Pflau T, Ruske G1 1Technical University of Munich, Germany, 2International Computer Science Institute, USA

Automatic Rhythm Modeling for Language Identification ........................................................................ 2539
Farinas J1, Pellegrino JF 1IRIT, France, 2DDL, France
For more information: http://www.irit.fr/ACTIVITES/EQ_ARTPS/

Session: E25 - Poster

Speech Recognition and Understanding: Confidence Measures and OOV

Confidence Measure (CM) Estimation for Large Vocabulary Speaker-Independent Continuous Speech Recognition System 2545
Zhang Y1, Lee R1, Madievski A2
1I.B.M. T.J.Watson Research Center, USA 2Motorola China Research Center, Beijing, China

Experimental Evaluation on Confidence of Agreement among Multiple Japanese LVCSR Models ..................... 2549
Kodama Y, Utsuro T, Nishizaki H, Nakagawa S Toyohashi University of Technology, Japan
For more information: http://www.cl.ics.tut.ac.jp/~utsuro/

Detection of Recognition Errors and Out of the Spelling Dictionary Names in a Spelled Name Recognizer for Spanish 2553
San-Seguido R, Macías-Guarasa J, Fereiros J, Martín P, Pardo J M Departamento de Ingeniería Electrónica, UPV, Spain
For more information: http://www.gth.die.upm.es
Non-english Abstract: ne1071.pdf

Use of acoustic prior information for confidence measure in ASR applications .............................................. 2557
Mengsugulog E, Riau C Faculté Polytechnique de Mons - TCTS lab, Belgium

Improving Performance of a Keyword Spotting System by Using a New Confidence Measure ........................ 2561
Ferrer L, Estienne C School of Engineering, University of Buenos Aires, Argentina

Word Level Confidence Measures Using N-Best Sub-Hypotheses Likelihood Ratio ........................................... 2565
Tan B T, Gu Y, Thomas T Vocals Ltd., UK

Confidence Based Lattice Segmentation and Minimum Bayes-Risk Decoding .............................................. 2569
Goel V, Kumar S, Byrne W2 1IBM, T.J.Watson Research Center, USA, 2Johns Hopkins University, USA

A Data Selection Strategy for Utterance Verification in Continuous Speech Recognition ............................. 2573
Jiang H, Soong P, Chin-Hui L Bell Labs, Lucent Technologies Inc., USA

Improved Speech Recognition Using Iterative Decoding Based on Confidence Measures ............................ 2577
Ogata J, Ariki Y Ryukoku University, Japan

Detection of OOV Words Using Generalized Word Models and a Semantic Class Language Model ................ 2581
Schaaf T University of Karlsruhe, Germany

Effects of OOV rates on Keyphrase Rejection Schemes ... 2585
Bouwman G, Sturm J, Boves L University of Nijmegen, the Netherlands

Session: E26 - Poster

Signal Analysis: Source Localisation and Beam Forming

A New Auditory Based Microphone Array and Objective Evaluation Using E-RASTI ..................................... 2591
Sanchez-Bote J-L, Gonzalez-Rodriguez J, Simon-Zorita D ATVS-DIAC-Universidad Politecnica de Madrid, Spain
For more information: http://www.atvs.diac.upm.es
Non-english Abstract: ne1043.pdf

Equivalence between Frequency Domain Blind Source Separation and Frequency Domain Adaptive Null Beamformers .... 2595
Araki S1, Makino S1, Mukai R1, Saruwatari H1 1NTT Communication Science Laboratories, Japan, 2Nara Institute of Science and Technology, Japan

Separation and Dereverberation Performance of Frequency Domain Blind Source Separation for Speech in a Reverberant Environment ................................................................. 2599
Mukai R, Araki S, Makino S NTT Communication Science Laboratories, Japan

Blind Source Separation for Speech Based on Fast-Convergence Algorithm with ICA and Beamforming .......... 2603
Saruwatari H, Kawamura T, Shikano K Nara Institute of Science and Technology, Japan
For more information: http://www.nara.ist-nara.ac.jp/~sawatari/

Noise Reduction Using Paired-microphones for both Far-field and Near-field Sound Sources ......................... 2607
Mizumachi M, Nakamura S ATR Spoken Language Translation Research Laboratories, Japan

Statistical Source Sound Identification in a Real Acoustic Environment for Robust Speech Recognition Using a Microphone Array .......................................................... 2611
Nishihara T3, Nakamura S1, Shikano K1 1ATR Spoken Language Translation Research Laboratories, Japan, 2Nara Institute of Science and Technology, Japan

Speech Enhancement and Source Separation based on Binaural Negative Beamforming ............................ 2615
Multiple source separation in the frequency domain using Negative Beamforming .............................................. 2019
Universidad Politécnica de Madrid, Spain

Planar Superdirective Microphone Arrays for Speech Acquisition in the Car .............................................. 2023
Martin R1, Petrovsky A2, Lotter T1
1Aachen University of Technology, Germany, 2Belarussian State University of Informatics and Radioelectronics, Republic of Belarus

Is Speech Data Clustered? - Statistical Analysis of Cepstral Features ...................................................... 2027
Kinnunen T, Fränti P
University of Joensuu, Finland

Using the Modulation Complex Wavelet Transform for Feature Extraction in Automatic Speech Recognition ............... 2039
Momomura Y1, Okada K1, Arai Y1, Kanedera N1, Murahara Y2
1Sophia Univ., Japan, 2Ishikawa National College of Technology, Japan

Separating Three Simultaneous Speeches with Two Microphones by Integrating Auditory and Visual Processing .................. 2043
Okuno H G, Nakadai K, Lourens T, Kitano H
Kitano Symbiotic Systems Project, IST, Japan
For more information: http://www.symbio.jst.go.jp/~okuno/

Signal Analysis: Speech Features and Modelling

A Time-Varying Complex AR Speech Analysis Based on GLS and ELS Method ...................................................... 2049
Funaki K
University of The Ryukus, Japan

Vocal Tract Normalization Equals Linear Transformation in Cepstral Space ...................................................... 2053
Pitz M, Molau S, Schlüter R, Ney H
RWTH Aachen - University of Technology, Germany

An Algorithm for finding Line Spectrum Frequencies of Added Speech Signals and its Application to Robust Speech Recognition ........................................................................................................... 2057
Yu A-T, Wang H-C
National Tsing Hua University, Taiwan

Improved entropic gain for speech signals analysis/synthesis based on an adaptive time-frequency segmentation scheme 2061
Gonon G1, Montresor S2, Baudry M1
1Laboratoire d’Informatique de l’Université de Maine, France, 2Laboratoire d’Acoustique de l’Université de Maine, France
For more information: http://www-ic2.univ-lemans.fr/~gonon/

Speech Recognition and Understanding: Kids, Toys and Emotions

Automatic Word Acquisition from Continuous Speech .................. 2067
Lucke H, Onote M
Sony Corporation, Japan
Non-english Abstract: ne1274.html

Why is automatic recognition of children’s speech difficult? .............. 2071
Li Q, Russell M
University of Birmingham, UK

Politeness and frustration language in child-machine interactions .......................................................................................... 2075
Arunachalam S, Gould D, Andersen E, Byrd D, Narayanan S
USC, USA

Speech Emotion Recognition Using Hidden Markov Models ............. 2079
Albino N, Asunción M, Antonio B, José B. M
Research Center TALP UPC, Spain
For more information: http://gps-tsc.upc.es/veu/

Applications: Media Applications

Speech Enhanced Remote Control for Media Terminal .................. 2085
Ibrahim A, Lundberg J, Johansson J
Linköpings Univ., Sweden

The development of a Portuguese version of a media watch system .............................................................................. 2089
Amaral R1, Langlois T1, Menedio H2, Neto F1, Souto N1, Trancoso F
1INESC ID/IPS, Portugal, 2INESC ID/IST, Portugal, 3INESC ID, Portugal
Non-english Abstract: ne1590.pdf

Classification of Video Genre using Audio .................................. 2093
Rouch M, Mason J
University of Wales Swansea, UK
For more information: http://galilee.swan.ac.uk

Prosody in Finger Braille and Teletext Receiver for Finger Braille .................................................................................. 2097
Horiiuchi Y, Ichikawa A
Chiba University, Japan

Speech Recognition and Understanding: Distributed Speech Recognition

Joint Channel Decoding - Viterbi Recognition For Wireless Applications .......................................................... 2073
Bernard A, Alwan A
UCLA, USA

MMSE-Based Channel Error Mitigation for Distributed Speech Recognition .......................................................... 2077
Peinado A M, Sanchez V, Segura J C, Perez-Cordoba J L
Universidad de Granada, Spain
Distributed Speech Recognition using Traditional and Hybrid Modeling Techniques ........................................ 2711
Stadermann J1, Meermeier R2, Rigoll G1
1University of Duisburg, Germany, 2SpeechWorks International, USA

Graceful Degradation of Speech Recognition Performance Over Lossy Packet Networks ..................................... 2715
Riskin E, Boulis C, Otterson S, Ostendorf M
University of Washington, USA

Session: E35 - Poster

Speech Recognition and Understanding: Prosody and Cross-Language in ASR

Experiments on Cross-language Acoustic Modeling ............... 2721
Schultz T, Wabel A
Carnegie Mellon University, USA
For more information:
http://www.cs.cmu.edu/~tanja/GlobalPhone/index.html

Crosslingual Speech Recognition with Multilingual Acoustic Models Based on Agglomerative and Tree-Based Triphone Clustering ................................................................. 2725
Zgank A1, Imperl B1, Johansen F1, Kacic Z2, Horvat B1
1University of Maribor, Slovenia, 2Telenor Research and Development, Norway

Comparing parameter tying methods for multilingual acoustic modelling ......................................................... 2729
Harju M1, Salmela P1, Leppanen J1, Viikki O1, Saarinen J1
1Tampere University of Technology, Finland, 2Nokia Research Center, Finland

Accent-Independent Universal HMM-Based Speech Recognizer for American, Australian and British English ........... 2733
Chengalvarayan R
Lucent Technologies Inc., USA

The Effect of Time Stress on Automatic Speech recognition Accuracy When Using Second Language .................. 2737
Chen F1, Sääv J1
1Swedish Center for Human Factors, Sweden, 2Linköping Univ., Sweden

The Effect of Pitch and Lexical Tone on Different Mandarin Speech Recognition Tasks .................................... 2741
Wong Y W1, Chang E2
1The Chinese University of Hong Kong, Hong Kong, P. R. China, 2Microsoft Research China, P. R. China

Acoustic Modeling of Foreign Words in a German Speech Recognition System ................................................. 2745
Stemmer G, Nöth E, Niemann H
University of Erlangen-Nürnberg, Germany

Semi-Automatic Grammatic Induction for Bi-directional English-Chinese Machine Translation ........................... 2749
Siu K-C, Meng H
The Chinese University of Hong Kong, Hong Kong, P. R. China

F0 Feature Extraction by Polynomial Regression Function for Monosyllabic Thai Tone Recognition ........................... 2753
Charnwuti P, Itatpukul S, Alkupatra V, Maneenoi E, Thathanon U, Thampanitchawong B
Chulalongkorn University, Thailand

The Use of Prosody in a Combined System for Punctuation Generation and Speech Recognition .................. 2757
Kim J-H, Woodland P C
Cambridge University, UK

Lexical Stress Modeling for Improved Speech Recognition of Spontaneous Telephone Speech in the Jupiter Domain .... 2761
Wang C, Seneff S
MIT Laboratory for Computer Science, USA

Modeling Auxiliary Information in Bayesian Network Based ASR ................................................................. 2765
Stephenson T A, Magimai Doss M, Bourlard H
Dalle Molle Institute for Perceptual Artificial Intelligence (IDIAP), Switzerland

A New Dynamic HMM Model for Speech Recognition .......... 2769
Chen F1, Chang E2
1Shanghai Jiao Tong University / Bell-lab Joint Lab, P. R. China, 2Microsoft Research China, P. R. China

Multi-Keyword Spotting of Telephone Speech Using Orthogonal Transform-Based SBR and RNN Prosodic Model .... 2773
Wang W-J, Chen S-H
National Chiao Tung University, Taiwan, ROC

Recognition of Slovenian Speech: Within and Cross-Language Experiments on Monophones using the SpeechDat(II) .... 2777
Iskra A1, Petek B1, Bredestad T2
1University of Ljubljana, Slovenia, 2Aalborg University, Denmark
For more information: http://isl.ttfex.uni-lj.si

Boiling down Prosody for the Classification of Boundaries and Accents in German and English ......................... 2781
Battiner A, Buckow J, Huber R, Warkne V, Nöth E, Niemann H
University of Erlangen-Nuremberg, Germany

Session: E36a - Poster

Education: Education and Training

JavaSpeakerRecognition - Interactive Workbench for Visualizing Speaker Recognition Concepts on the WWW ........ 2787
Drygajlo A, Garcia Molina G
EPFL, Switzerland

Prototype of a Vocal-Tract Model for Vowel Production Designed for Education in Speech Science .................. 2791
Arai T, Usuki N, Murahara Y
Sophia University, Japan

A tool for automatic feedback on phonemic transcription .... 2795
Cooke M1, Garcia Lecumberri M L2, Maidment J3
1University of Sheffield, UK, 2University of the Basque Country, Spain, 3University College London, UK
Non-english Abstract: nei527.html

Speech Lab in a Box: A Mandarin Speech Toolbox to Jumpstart Speech Related Research ................................. 2799
Chang E, Shi Y, Zhou J, Huang C
Microsoft Research China, P. R. China

Relating PhonePass Scores Overall Scores to the Council of Europe Framework Level Descriptors ......................... 2803
de Jong J H A L1, Bernstein J2
1Language Testing Services, the Netherlands, 2Ordinate Corporation, USA
For more information: http://www.ordinate.com
A Multilingual, Multimodal, Speech Training System, SPECO 2807
Vicsi K1, Roach P2, Oster A-M3, Kacic Z4, Csatári F5, Sfakianaki A6, Veronik R7, Gordos G8
1Budapest University of Technology and Economics, Hungary, 2University of Reading, UK, 3Kungl. Tekniska Högskolan, Sweden, 4University of Maribor, Slovenia
For more information: http://luna.ttt.bme.hu/speech/

Instantaneous Estimation of Accentuation Habits for Japanese Students to Learn English Pronunciation 2811
Nakamura N1, Minematsu N2, Nakagawa S3
1Toyohashi University of Technology, Japan, 2University of Tokyo, Japan

Automatic Construction of CALL System from TV News Program with Captions 2815
Tanaka T, Mori K, Kobayashi S, Nakagawa S
Toyohashi University of Technology, Japan
Non-english Abstract: ne1321.pdf

Speaker Recognition: Features and Robustness

Pitch-Dependent GMMs for Text-Independent Speaker Recognition Systems 2821
Arcienega M, Drygajlo A
EPFL, Switzerland

Towards combining pitch and MFCC for speaker recognition systems 2825
Ezzaidi H1, Rouat J1, O'Shaughnessy 2
1DSA, ERMETIS, Université du Québec à Chicoutimi, Canada, 2INRS-Telecommunications, Université du Québec, Canada

Formant-Broadened CMS Using Peak-Picking in LOG Spectrum 2829
Kim Y-J, Jung H-K, Chung J-H
Inha Univ., Korea

Improvements in the speaker identification rate using feature-sets 2833
Mashao D J, Baloyi N T
University of Cape Town, South Africa

Minimum Classification Error Training for Speaker Identification Using Gaussian Mixture Models Based on Multi-Space Probability Distribution 2837
Miyajima C, Tokuda K, Kitamura T
Nagoya Institute of Technology, Japan

Speaker Recognition based on Feature Space Trace 2841
Wu Y, Li Z
Shanghai Jiaotong University, P. R. China

Additive and Convolutional Noise Canceling in Speaker Verification Using a Stochastic Weighted Viterbi Algorithm 2845
Yoma N B, Villar M
University of Chile, Chile
For more information: http://www.die.uchile.cl

A Multi-SNR Subband Model for Speaker Identification under Noisy Environments 2849
Yoshida K, Takagi K, Ozeki K