

**PROCEEDINGS OF THE
EIGHTH AUSTRALIAN
INTERNATIONAL CONFERENCE ON
SPEECH SCIENCE AND TECHNOLOGY**

**Canberra
5-7 December 2000**

Edited by
Michael Barlow
University of NSW/ADFA

With Assistance from
Philip Rose
Australian National University

Australian Speech Science and Technology Association (Inc)
Canberra 2000

© Australian Speech Science and Technology Association (Inc) 2000

This work is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced by any process without written permission from the Australian Speech Science and Technology Association (Inc). Inquiries should be directed to the Secretary, Australian Speech Science and Technology Association, GPO Box 143, Canberra, ACT 2601, Australia.

ISSN 1039-0227

ISBN 0 9588579 8 9

Cover Picture: Photograph of Parliament House, by David Paterson. M. Photog. Mosaic pavement, Parliament House forecourt. Designed by Michael Nelson Tjakamarra and fabricated by William McIntosh, with Aldo Rossi and Franco Colossi.

FOREWORD

SST-2000 is the eighth in the biennial series of Australian International Conferences on Speech Science and Technology. It was held at the Australian National University from the 5th to the 7th of December 2000. The conference series is organised by ASSTA (the Australian Speech Science & Technology Association) Inc.

SST-2000 sees the series "come full circle" by returning to its city of origin Canberra, after having been hosted by most of the capital cities and centres of speech research in Australia: SST-86, the first, was held in Canberra; with SST-88 in Sydney; SST-90 in Melbourne; SST-92 in Brisbane; SST-94 in Perth and SST-96 in Adelaide. SST-98 was a special-case, run as an adjunct to the larger ICSLP-98 and specially targeted at postgraduate students working in the speech field. It is illuminating to look back through previous proceedings (and memories of those conferences) with an eye both to the indelible and ephemeral qualities that have marked the SST series.

Clearly, the goal of the SST series, since its inception in 1986, has been to serve as a forum at which Australian and international researchers working in the multidisciplinary fields of speech science and speech technology may come together to freely share ideas and interactions. I believe that to-date there can be little doubt that the series has succeeded in that role. That said though, there is no room for complacency in the organisation or vision that drives the SST series, if the series is to maintain its key and central role.

SST-2000 has seen a number of innovations aimed at strengthening the conference. Perhaps chief amongst these has been the far more rigorous review process. For the first time full-paper submission for review has been offered as an option, with authors free to choose from that or the traditional extended abstract. A comprehensive international and national review team has been employed, and this process has been facilitated by the use of electronic submission and review: greatly speeding and standardising the process. Indeed Internet technology has proved vital as an enabler for the conference, with the web site serving as a clearing-house for submissions and reviews as well as the other detailed information concerning conduct of the conference (programme, presentation details, awards, etc.).

In all there are eighty papers in the proceedings, organised alphabetically by major topic, and within topic alphabetically by first-author's surname. It is pleasing to see the depth and diversity of Australian (and international) research that these papers represent. I would like to take the opportunity to publicly congratulate Jennifer Elliot, the winner of the ASSTA New Researcher Award, as well as to thank my fellow organising committee members and the conference attendees for their efforts and involvement in SST-2000.

Michael Barlow
Chairperson SST-2000

SPONSORS

Cochlear Pty. Ltd.
School of Computer Science, UNSW/ADFA
Appen Speech Technology Pty. Ltd.
School of Language Studies, ANU

ORGANISING COMMITTEE

Chairperson	Michael Barlow	The University of NSW / ADFA
Secretary	Michael Barlow	The University of NSW / ADFA
Treasurer	Shunichi Ishihara	The Australian National University
Technical Program & Proceedings	Michael Barlow Philip Rose	The University of NSW / ADFA The Australian National University
Publicity	Roland Goecke	The Australian National University
Tutorial Day	Frantz Clermont	The University of NSW / ADFA
Sponsorship	Michael Wagner	The University of Canberra
ASSTA Liaison	Bruce Millar	The Australian National University
Web Page Maintainer	Michael Barlow	The University of NSW / ADFA

Corporate Members of ASSTA Inc.

ASSTA Inc. is pleased to acknowledge the following companies as its Corporate Members during 2000.

Appen Pty Ltd
Cochlear Ltd
CRC for Cochlear Implant and Hearing Aid Innovation
Motorola Australian Research Centre
Syrinx Speech Systems Pty Ltd

REVIEW PROCESS

All submissions (whether full-paper or extended abstract) to SST-2000 were blind peer reviewed. The process was overseen by a Technical Programme Committee of two members who coordinated the entire process. Reviewers were drawn from two boards: a National (Australian) Review Board and an International Review Board.

A sub-committee of four members was also formed to select a winner for the ASSTA New Researcher Award (NRA).

Membership of those boards and committees are found below. All members of the Review Boards listed contributed to the review process.

TECHNICAL PROGRAM COMMITTEE

Michael Barlow

Philip Rose

ASSTA NEW RESEARCHER AWARD COMMITTEE

Michael Barlow

Bruce Millar

Frantz Clermont

Philip Rose

INTERNATIONAL REVIEW BOARD

William Ainsworth

Christoph Draxler

Cheolwoo Jo

Toomas Altsaar

Kjell Elenius

Beth Logan

Frederic Berthommier

Mark Forsyth

Doug O'Shaughnessy

Steven Bird

Sadaoki Furui

Tomoko Matsui

Matts Blomberg

Phil Green

Parham Mokhtari

Elizabeth Botha

Shuichi Itahashi

Tuan Pham

Nick Campbell

Jagath-Kumara

Justus Roux

Martin Cooke

Chris James

NATIONAL REVIEW BOARD

Eliathamby Ambikairajah

Jan van Doorn

Bruce Millar

Peter Blamey

Helen Fraser

Jenni Oates

Denis Burnham

David Grayden

Kuldip Paliwal

Frantz Clermont

John Ingram

Roberto Togneri

Chris DeSilva

Robin King

Michael Wagner

SUBMISSIONS ACCEPTED ON THE BASIS OF FULL-PAPER REVIEW

Listed below are those submissions to SST-2000 which were **reviewed and accepted on the basis of the full-paper**. This separate list is provided because authors submitting to the conference were given the choice of either submitting a full paper, or extended abstract, for review.

All submissions to the conference were blind (all author details excised from the submission) peer-reviewed by members of a National and International Review Board. The process was overseen by the Technical Programme Committee who assigned reviewers for each submission received, together with collating the review results.

ACCEPTED FULL PAPERS

Listed below, and ordered on the basis of first author's surname, are all submissions accepted for presentation at SST-2000 that were submitted and reviewed as full papers.

Ahmed B., and Holmes W.H.

Objective and subjective performance measures for voice activity detectors

Athaudage C., Bradley A.B., and Lech M.

Efficient compression of melp spectral parameters using optimized temporal decomposition

Barlow M., Watson B., Tsoi A-C., and Downs T.

A-priori selection of cohort sets for a speaker verification system: issues and insights

Barlow M., and Wagner M.

Perceptions of identity, gender and idiolect in prosodically altered speech using a composite model approach

Blamey, P., James C., and Martin L.

Sound separation with a cochlear implant and a hearing aid in opposite ears

Burnham D., Ciocca V., Lauw C., Lau S., and Stokes S.

Perception of visual information for cantonese tones

Carrera-Sabaté J., Fernández-Planas A.M., Matas-Crespo J., Ortega-Escandell A.

Differences in vowel quality in two Catalan dialects. Data from MDS.

Carson-Berndsen J., and Walsh M.

Interpreting multilinear representations in speech

Cassidy S., Welby P., McGory J., and Beckman M.

Testing the adequacy of query languages against annotated spoken dialog

Dines J., Sridharan S., and Moody M.

Compression of speech for mass storage using speech recognition and text-to-speech synthesis

van Doorn J.

Does artificially increased speech rate help?

Elliott J.

Auditory and F-pattern variations in Australian okay: a forensic phonetic investigation

Elliott J.

Comparing the acoustic properties of normal and shouted speech: a study in forensic phonetics

- Epps J., and Holmes W.H.*
Wideband speech coding at narrowband bit rates
- Flax M., Ambikairajah E., and Holmes W.H.*
Improved auditory masking models
- Goecke R., Tran Q.N., Millar B., Zelinsky A., and Robert-Ribes J.*
Validation of an automatic lip-tracking algorithm and design of a database for audio-video speech processing
- Ishihara S.*
Continuous linguistic tonetic representation using polynomial residuals
- Ishihara S.*
Linguistic-tonetic differences in target-tone realisation: standard vs. Kagoshima Japanese
- Ishihara S.*
The exponential nature of F0 target tone interpolation
- Kim H-G., Obermayer K., Bode M., and Ruwisch D.*
A 1.6 kbps speech codec using spectral vector quantization of differential features
- Kim H-G., Obermayer K., Bode M., and Ruwisch D.*
Real-time noise cancelling based on spectral minimum detection and diffusive gain factors
- Kinoshita Y.*
Effective F2 as parameter in Japanese forensic speaker identification
- Kinoshita Y., and Maironald J.*
Statistical quantification of differential vowel comparability in forensic phonetic samples
- Kovtonyuk A., Kalyuzhny, and Semenov*
Adaptive Kalman filtering of speech signals, based on a block model in the state space and vector quantization of auto regressive features
- Lucey S., Sridharan S., and Chandran V.*
A improvement of automatic speech reading using a intensity to contour stochastic transformation
- McCowan I., Moore D., and Sridharan S.*
Speech enhancement using near-field superdirectivity with an adaptive sidelobe canceler & postfilter
- Martin L, Blamey P., James C., Galvin K., and Macfarlane D.*
Adaptive dynamic range optimisation for hearing aids
- Mason M., Sridharan S., and Chandran V.*
A comparison of two hybrid audio coding structures incorporating discrete wavelet transforms
- Millar B.*
Prospects for speech technology in the Oceania region
- Ming L., Junkawitsch J., and Yu T.*
An incremental approach to selection of well balanced corpus
- Myers S., Pelecanos J., and Sridharan S.*
Two speaker detection by dual gaussian mixture modelling
- Pelecanos J., Myers S., and Sridharan S.*
Rapid channel compensation for one and two speaker detection in the NIST 2000 speaker recognition evaluation

Rose P.

Hong Kong Cantonese citation tone acoustics: a linguistic tonetic study

Rose P.

Wenzhou dialect disyllabic lexical tone sandhi with first syllable entering tones

Tsurutani C. , and Ingram J.

Perception of moraic timing by English learners of Japanese

Veprek P., and Bradley A.B.

Hierarchical speech compression for storage - a two-level approach

Wark T., Sridharan S., and Chandran V.

A comparison of static and dynamic classifier performance for multi-modal speaker verification

Wojdel J. and Rothkrantz L.

Silence detection and vowel/consonant discrimination in video sequences

Wong E., Pelecanos J., Myers S. and Sridharan S.

Language identification using efficient gaussian mixture model analysis

CONTENTS

KEYNOTE ADDRESSES

Beckman M. and Pierrehumbert J.

Positions, probabilities, and levels of categorisation 2

Furui S.

Steps towards flexible speech recognition – recent progress at
Tokyo Institute of Technology - 19

DIALOGUE SYSTEMS

Kiriyama S., Hirose K., and Minematsu N.

Development and evaluation of a spoken dialogue system for academic document
retrieval with a focus on reply generation 32

Wang X.F., and Du L.M.

The design of a Chinese spoken dialogue system engine 38

LANGUAGE ACQUISITION

Czigler P., Van Doorn J. and Sullivan K.

An acoustic study of the development of word-initial /sP/ consonant clusters in the
speech of a Swedish child aged 1;11-2;5 years 44

Kim C-W., and Ingram J.

Perception of Korean back vowels by Australian English and Japanese speaking
adult learners 50

Tsukada K.

Some acoustic characteristics of Australian English /aI/ and Japanese /ai/ in native
and non-native speech production 56

Tsurutani C., and Ingram J.

Perception of moraic timing by English learners of Japanese 62

LANGUAGE IDENTIFICATION

Sheikhzadegan J., and Roohani M.

Automatic spoken language identification based on ANN using fundamental frequency
and relative changes in spectrum 70

Willmore J., Price R., and Roberts W.

Comparing gaussian and neural network modelling approaches to automatic language
identification of speech 74

Wong E., Pelecanos J., Myers S., and Sridharan S.

Language identification using efficient gaussian mixture model analysis 78

MULTIMODAL SPEECH

Burnham D., Ciocca V., Lauw C., Lau S., and Stokes S.

Perception of visual information for Cantonese tones 86

Goecke R., Tran Q.N., Millar B., Zelinsky A., and Robert-Ribes J.

Validation of an automatic lip-tracking algorithm and design of a database for
audio-video speech processing 92

<i>Lucey S., Sridharan S., and Chandran V.</i> An improvement of automatic speech reading using an intensity to contour stochastic transformation	98
<i>Wojdel J. and Rothkrantz L.</i> Silence detection and vowel/consonant discrimination in video sequences	104
PHONETICS - ACOUSTIC	
<i>Barlow M., and Clermont F.</i> A parametric model of Australian English vowels in formant space	112
<i>Barlow M., and Clermont F.</i> Seeing is believing: beyond a static 2D-view of formant space for research and education	118
<i>Carrera-Sabaté J., Fernández-Planas A.M., Matas-Crespo J., Ortega-Escandell A.</i> Differences in vowel quality in two Catalan dialects. Data from MDS.	124
<i>Harrington J., Palethorpe S., and Watson C.</i> Vowel change in Received Pronunciation: evidence from the Queen's English	130
<i>Onaka, A., and Watson C.</i> Acoustic comparison of child and adult fricatives	134
<i>Thompson N.</i> Wuyi citation tone acoustics: problems for tonological representation	140
PHONETICS - FORENSIC	
<i>Elliott J.</i> Auditory and f-pattern variations in Australian okay: a forensic phonetic investigation	148
<i>Elliott J.</i> Comparing the acoustic properties of normal and shouted speech: a study in forensic Phonetics	154
<i>Kinoshita Y.</i> Effective F2 as a parameter in Japanese forensic speaker identification	160
<i>Kinoshita Y., and Mairdonald J.</i> Statistical quantification of differential vowel comparability in forensic phonetic samples	166
<i>Rose P., and Clermont F.</i> Comparative performance of cepstrum- and formant-based analysis on similar-sounding speakers for forensic speaker identification	172
PHONETICS - LINGUISTIC	
<i>Fraser H.</i> Phonetics, phonology and the teaching of pronunciation – a new cd-rom for ESL learners, and its rationale	180
<i>Gharavian D., Sheikhzadeh H., and Ahadi S.M.</i> An experimental multi-speaker study on farsi phoneme duration rules using automatic Alignment	186
<i>Ishihara S.</i> Linguistic-tonetic differences in target-tone realisation: standard vs. Kagoshima Japanese	192

Rose P.
Hong Kong Cantonese citation tone acoustics: a linguistic tonetic study 198

PROSODY

D'Imperio M., Terken J., and Pitermann M.
Perceived tone "targets" and pitch accent identification in Italian 206

Hansson P.
Focal accentuation and boundary perception 212

Ishihara S.
Continuous linguistic tonetic representation using polynomial residuals 218

Ishihara S.
The exponential nature of F0 target tone interpolation 224

Rose P.
Wenzhou dialect disyllabic lexical tone Sandhi with first syllable entering tones 230

Yim H-S.
An analysis of Korean intonation in declarative and propositive sentence types 236

SIGNALS PROCESSING & FEATURE ANALYSIS

Flax M., Ambikairajah E., and Holmes W.H.
Improved auditory masking models 244

Kim H-G., Obermayer K., Bode M., and Ruwisch D.
Real-time noise cancelling based on spectral minimum detection and diffusive gain factors 250

Kim H.S., and Holmes W.H.
Nonparametric peak feature extraction and its application to speech signals 256

Kovtonyuk A., Kalyuzhny A., and Semenov V.
Adaptive Kalman filtering of speech signals, based on a block model in the state space and vector quantization of auto regressive features 262

McCowan I., Moore D., and Sridhran S.
Speech enhancement using near-field superdirectivity with an adaptive sidelobe canceler & postfilter 268

Phooi S.K., Man Z., and Wu H.R.
A new approach in designing an adaptive lattice predictor for nonlinear and nonstationary speech signals in ADPCM using Lyapunov theory 274

Thorpe C.W., and Watson C.
Vowel identification in singing at high pitch 280

SPEAKER RECOGNITION

Barlow M., Watson B., Tsoi A-C., and Downs T.
A-priori selection of cohort sets for a speaker verification system: issues and insights 288

Demenko G.
Analysis of suprasegmental features for speaker verification 294

<i>Myers S., Pelecanos J., and Sridharan S.</i> Two speaker detection by dual gaussian mixture modelling	300
<i>Pelecanos J., Myers S., and Sridharan S.</i> Rapid channel compensation for one and two speaker detection in the NIST 2000 speaker recognition evaluation	306
<i>Sanderson C., and Paliwal K.</i> Training method of a piecewise linear classifier for an adaptive multi-modal person verification system	312
<i>Wark T., Sridharan S., and Chandran V.</i> A comparison of static and dynamic classifier performance for multi-modal speaker verification	318
<i>Wildermoth B., and Paliwal K.</i> Use of voicing and pitch information for speaker recognition	324
SPEAKER CHARACTERISTICS	
<i>Barlow M., and Wagner M.</i> Perceptions of identity, gender and idiolect in prosodically altered speech using a composite model approach	330
<i>Sullivan K., Bayard D., Weatherall A., Gallois C., Schlichting F., and Pittam J.</i> Does media exposure to an accent impact upon the estimation of the age of speakers?	336
<i>Zetterholm E.</i> The significance of phonetics in voice imitation	342
SPEECH AIDS & DISORDERS	
<i>van Doorn J.</i> Does artificially increased speech rate help?	350
<i>Grayden D., and Clark G.</i> The effect of rate of stimulation of the auditory nerve on phoneme recognition	356
<i>Griffin S., Wilson L., and Clark E.</i> Speech pathology applications of speech recognition technology	362
<i>Jo C., Kim D., Baek M., and Wang S.</i> On predicting patient's voice after surgical operation	367
<i>Martin L, Blamey P., James C., Galvin K., and Macfarlane D.</i> Adaptive dynamic range optimisation for hearing aids	373
SPEECH CODING	
<i>Ahmed B., and Holmes W.H.</i> Objective and subjective performance measures for voice activity detectors	380
<i>Athaudage C., Bradley A.B., and Lech M.</i> Efficient compression of melp spectral parameters using optimized temporal decomposition	386
<i>Dines J., Sridharan S., and Moody M.</i> Compression of speech for mass storage using speech recognition and text-to-speech synthesis	392

<i>Epps J., and Holmes W.H.</i> Wideband speech coding at narrowband bit rates	398
<i>Kim H-G., Obermayer K., Bode M., and Ruwisch D.</i> A 1.6 kbps speech codec using spectral vector quantization of differential features	404
<i>Mason M., Sridharan S., and Chandran V.</i> A comparison of two hybrid audio coding structures incorporating discrete wavelet transforms	410
<i>Ritz C., and Burnett I.</i> Split temporal decomposition and quantisation	416
<i>Veprek P., and Bradley A.B.</i> Hierarchical speech compression for storage - a two-level approach	421
SPEECH DATABASES	
<i>Cassidy S., Welby P., McGory J., and Beckman M.</i> Testing the adequacy of query languages against annotated spoken dialog	428
<i>Millar B.</i> Prospects for speech technology in the Oceania region	434
<i>Ming L., Junkawitsch J., and Yu T.</i> An incremental approach to selection of well balanced corpus	440
SPEECH PERCEPTION	
<i>Behne D., Czigler P., and Sullivan K.</i> Perception of Swedish vowel quantity: tracing stages of development	446
<i>Blamey, P., James C., and Martin L.</i> Sound separation with a cochlear implant and a hearing aid in opposite ears	452
<i>Tyler M.D., and Burnham D.</i> Orthographic influences on initial phoneme deletion tasks	458
SPEECH PHYSIOLOGY	
<i>Barlow M., Clermont F., and Mokhtari P.</i> From acoustics of speech to a 3D vocal-tract: towards a plausible model with real-time constraints	466
<i>Carson-Berndsen J., and Walsh M.</i> Interpreting multilinear representations in speech	472
<i>Mokhtari P., and Clermont F.</i> New perspectives on linear-prediction modelling of the vocal- tract: uniqueness, formant-dependence and shape-parameterisation	478
<i>Robinson R.</i> Articulograph interface	484
SPEECH RECOGNITION	
<i>Ahadi M.</i> Reduced context sensitivity in Persian speech recognition via syllable modeling	492

<i>Bain K., and Paez, D.</i> Speech recognition in lecture theatres - liberated learning project an innovation to improve access to higher education using speech recognition technology	498
<i>Feng J., and Du L.M.</i> An improved architecture for word verification	504
<i>Jeon H., Ryu C-S., Kim J-I., and Koo M-W.</i> A speech-operated railway information & reservation service with multistage dialogue	510
AUTHOR INDEX	519