

*Australian Speech Science
and Technology Association
(Incorporated)*
G.P.O. Box 143, Canberra

ASSTA Newsletter

April 2002

Editorial

Welcome to the March 2002 issue of the ASSTA Newsletter. You will notice a series of new faces in this issue, starting with the one to the right! My name is Michael Tyler, and I am your new Newsletter Editor.

First, I would like to take the opportunity to thank Chris James for his time as newsletter editor. Chris brought to the newsletter new ideas, including the 5th Column, which is a section where you can air your views about ASSTA, or Speech Science and Technology in Australia. This column will continue, and I encourage you all to send me your views, no matter how controversial. I have started the ball rolling in this issue, with a piece on multidisciplinary interactions.

The other new faces are those of our new ASSTA executive, with



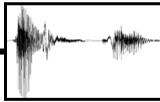
a new President, Treasurer, and Ordinary Member. Given that there has been such a change in the executive, we have provided a short introduction for each member, which you can find on page 6.

I look forward to my role as editor. Please feel free to contact me with material and ideas.

Michael Tyler

Contents

<u>President's Report</u>	<u>3</u>
<u>Research Initiatives</u>	<u>5</u>
<u>The New Executive</u>	<u>6</u>
<u>Forensic Speaker Identification</u>	<u>8,9</u>
<u>Speech Resources</u>	<u>11</u>
<u>Conference Log</u>	<u>12</u>
<u>The 5th Column</u>	<u>15</u>
<u>SST-2002</u>	<u>16</u>



ASSTA Information

ASSTA Executive

President: Denis Burnham
d.burnham@uws.edu.au
Fax: (02) 9772 6736

Secretary: Steve Cassidy
steve.cassidy@mq.edu.au

Treasurer: Shunichi Ishihara
Shunichi.Ishihara@anu.edu.au

Member: David Grayden
D.Grayden@medoto.unimelb.edu.au
Fax: (03) 9283 7518

Member: Johanna Barry
J.Barry@medoto.unimelb.edu.au

ASSTA Newsletter Team

Editor: Michael Tyler
m.tyler@uws.edu.au
Fax: (02) 9772 6736

Sub-ed.(conferences): Marija Tabain
marija@maccs.mq.edu.au

Sub-ed.(reviews/abstracts): Johanna Barry
J.Barry@medoto.unimelb.edu.au

Contributions

Copy should be sent to the editor via Email as plain text or as an attachment in Rich Text Format (RTF).

Send notices or reports on conferences to [Marija Tabain](#). Thesis abstracts and book reviews should be sent to [Johanna Barry](#).

The new address for hardcopy is:

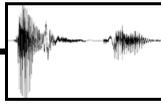
Michael Tyler
MARCS, University of Western Sydney
Bankstown Campus, Building 24
Locked Bag 1797, South Penrith D.C.
South Penrith, NSW, 1797.

Contributors to this issue: Denis Burnham, Roberto Togneri, Marija Tabain, Yuko Kinoshita, Phil Rose.

ASSTA recognises the assistance of the *Macarthur Auditory Research Centre Sydney (MARCS)* and the *University of Western Sydney* in the production and distribution of this newsletter.

Please visit ASSTA on the Web:

<http://www.assta.org>



President's Report

Denis Burnham



Let me begin with a tribute to the outgoing President of ASSTA, Bruce Millar. A stalwart of speech science in Australia, Bruce was intimately involved in the formation of ASSTA, and has been on the Executive for the past 13 years and President for the latter six. Bruce has played a pioneering role in the establishment of Speech Science and Technology in Australia. On behalf of all ASSTA members, thank you, Bruce, for your time on the Executive and as President. With your great leadership and competence you have left an imposing task for us who follow you.

In the past few years, our international colleagues have become more accessible through a combination of the decreased cost of travel, increased availability and accessibility of international conferences, electronic mail, and the lightening-fast transfer of text, visual,

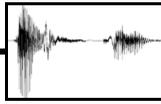
and auditory data. Correspondingly, the role of national scholarly and professional bodies is undergoing some scrutiny. Can local conferences survive? What do and can national bodies offer?

Here I will take up but one issue regarding this, the vital role that ASSTA can make to foster and promote collaborative research and development within Australia.

There is a depth of knowledge and expertise in Speech Science and Technology in Australia. ASSTA was actually formed before one of our international counterparts, ISCA, the International Speech Communication Association (née ESCA). This depth can be seen at the SST (Speech Science and Technology) conferences, which provide an ideal opportunity to get together, talk science, see new students displaying their wares, and set up collaborations. However, SST comes but once every even year, and there is often the thought that as the peak body of Speech Science and Technology in Australia, ASSTA should more consistently and actively facilitate communication and collaboration between our colleagues.

I write this in a hotel room on the eve of attending two successive ISCA

(Continued on page 4)



(Continued from page 3)

workshops in the south of France (well someone has to do it!), on different aspects of speech science. ISCA can mount these workshops for two main reasons: because in Europe, and in nearby North America, there is the population base to support such workshops; and because they have funds.

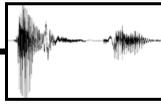
Thanks to the hard work of a whole lot of people, including especially our outgoing President, we also now have discretionary funds available – profit from running ICSLP, the International Conference on Spoken Language Processing, in Sydney in 1998. Accordingly, the ASSTA executive has recently launched "ASSTA Research Events", a scheme by which ASSTA members can apply for money to conduct seminars, workshops, initial meetings for new collaborations, or anything that the members can think of which will promote research and scholarship in Speech Science and Technology. This is focussed on Australian research, but may of course involve international colleagues. Such funds provided by ASSTA will also be useful in attracting further funds from other sources in the current climate of sponsorship and matching grants.

We favour a grass roots approach in this scheme; we want members to set the agenda for these events. Just to underline this - ASSTA has money to give to its members to promote research. Please ask for it! Other ASSTA initiatives include the 'PhD of the Year'

award, and the ASSTA study awards. Go to the website, look at the guidelines, and see the descriptions regarding these schemes on the next page of this newsletter.

We have entered an exciting era for Speech Science and Technology: there are advances and emerging interests in various areas such as ASR, TTS systems, avatars and talking heads, the role of visual information in speech, emotion in speech, cross-language and multi-language studies which engage the full range of our contributing disciplines, from the more technical machine end of the spectrum, to the more human information processing end. Coupled with this we find ourselves in an interesting period regarding the relationship between national and international scholarly bodies, and regarding the diversity of opportunities for research funding. I will talk more of these issues in later reports. For now, let me say that I look forward to serving you as President of ASSTA and in these interesting times. I particularly look forward to seeing ASSTA research events funded.

Denis Burnham



ASSTA Research Initiatives

Each of the current initiatives will be described briefly below, but the full eligibility criteria, and the application procedures, are described on the ASSTA website: <http://www.assta.org/initiatives/>. These schemes are for you, so please apply!

PhD of the Year Award

It is our great pleasure to announce that **Cecile Pereira** has been awarded this year's ASSTA PhD of the Year award for thesis entitled "Perception and Expression of Emotion in Speech". Congratulations Cecile!

The PhD of the year award is open to applicants who have graduated from a Australian PhD course within the last 15 months. It is a \$1000 prize for the best PhD submitted by ASSTA members.

New Researcher Award

This award provides airfare, registration, college accommodation and a conference dinner ticket to the SST conference. Applicants must be post-graduate students or early researchers who have graduated from a postgraduate degree in the past 5 years.

Applications are due by the deadline

for abstracts for the SST conference. It should be noted that the full conference paper must be submitted with the application.

PhD Study Awards

Up to six PhD study awards, of \$500 each, are offered annually to students who have been enrolled in a full-time PhD course for between 6 and 18 months.

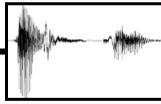
Applications for these awards close on May 3, 2002.

Research Event Awards

As mentioned in the President's report, these are a new initiative, designed to encourage collaborative interactions between members and other researchers between SST conferences. There are three classes of events:

- (1) seminars centred around a particular theme;
- (2) groundwork meetings for setting up or strengthening research collaborations;
- (3) workshops on particular issues of note, new theories, procedures, approaches, etc.

The closing date for applications is May 31, 2002.



The New ASSTA Executive

President: Denis Burnham

Denis Burnham is a Professor of Psychology and the Director of the Macarthur Auditory Research Centre Sydney (MARCS), at the University of Western Sydney, and a foundation member of ASSTA. For the last 20 years, Denis has conducted research into human speech perception abilities, in both adults and children.

His current research interests include cross-language speech perception, auditory visual speech perception, auditory and auditory-visual perception of tone, speech perception development and its relation to reading ability, accessibility and availability of captions for the Deaf and Hearing-Impaired, and the evaluation of TTS systems.

Denis enjoys worldwide collaborations with researchers in Thailand, Japan and Hong Kong.

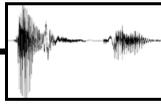
Secretary: Steve Cassidy



Steve Cassidy is a Computer Scientist who has worked in various areas relating to language and cognition over the last 15 years. He completed a PhD in Wellington, New Zealand on computer models of reading development and then moved to Macquarie University, Sydney to work in the Speech Hearing and Language Research Centre (SHLRC). At SHLRC he worked on applying statistical models to acoustic phonetics problems and on the development of the Emu Speech Database System. His work on Emu has led to an involvement with groups in the US and Europe who are aiming to define standards for Linguistic annotation. Steve is now working in the Computing Department at Macquarie where he is pursuing research in Linguistic annotation and speech technology systems.

***This ASSTA Newsletter
is on the Net!***

[http://www.assta.org/
news/2002_1/](http://www.assta.org/news/2002_1/)



Treasurer: Shunichi Ishihara



Shunichi is currently teaching Japanese Language and Linguistics at the Japan Centre of the Australian National University (ANU), while completing a PhD thesis in experimental phonetics as a part time student, also at ANU. This is Shunichi's first year on the ASSTA executive.

ASSTA executive. She is interested in speech development in profoundly hearing-impaired children. Her research includes, among other things, a longitudinal study of speech development in Cantonese-speaking children who have received a cochlear implant.

Member: David Grayden

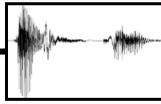


David Grayden has been working as a Research Fellow at the Bionic Ear Institute in Melbourne since 1997. His main research involves careful examination of phoneme confusions made by cochlear implant users, with the view to designing strategies to improve perception. He has also been working in a joint research project with Telstra Research Laboratories where direct readings are obtained from animal brain-stems to learn about the role of this structure in speech encoding. This knowledge will be used to improve the cochlear implant and for the improvement of front-end processing for automatic speech recognition.

Member: Johanna Barry



Johanna Barry is a PhD student from the University of Melbourne, studying in the departments of Otolaryngology and Linguistics & Applied Linguistics. This is also Johanna's first year on the



Ph.D. Thesis Abstract

Reviews compiled and edited by Johanna Barry



Ph.D. Thesis: *Testing realistic forensic speaker identification in Japanese: A likelihood ratio-based approach using formants.*

Author: Yuko Kinoshita, The Australian National University

This thesis investigates to what extent it is possible to perform forensic speaker identification under three important conditions of realism:

- With forensically realistic data – i.e. natural and non-contemporaneous speech.
- With ‘traditional’ (i.e. non-automatic) acoustic parameters of formant centre-frequency.
- With a Bayesian Likelihood Ratio approach as a discriminant measure.

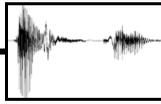
Natural speech data were elicited from 13 male speakers of Standard Japanese on two occasions separated by about a fortnight. A map task was used which was designed to contain words with five tokens of each of the short pitch-accented vowels /i e a o u/ e.g. *nemoto name*, *terebi television*, *sushiya sushi shop*, *jinja shrine*. In addition, a set of several tokens of *moshimoshi* (the Japanese word for *hello* over the phone) was elicited. The task was repeated

twice in each recording session, giving four sets of data overall. The centre-frequencies of the vowels’ first four formants, and the F-pattern in *moshi-moshi* were extracted by LPC analysis using CSL, and then subjected to further statistical processing.

An ANOVA was carried out to determine what segment / formant combinations were more promising as speaker identification parameters from the point of view of magnitude of F-ratio. On the basis of high F-ratio, and to conform to forensic realism, three values were selected from the individual vowels, and three from the set of *moshimoshi*, to be used as discriminant parameters. For the vowels, these were F2 and F3 in /e/ and F2 in /i/. For *moshimoshi*, these were F3 in /m2/, /o1/ and /s2/.

A discriminant test was performed using an overall likelihood ratio as a discriminant function. This was calculated by first finding the likelihood ratio for each of the six parameters, checking that they were not significantly correlated, and then taking their product to derive the overall likelihood ratio. Means and standard deviations for the reference population in the formula were estimated using a cross-validation (leave-one-out) technique. Values of

(Continued on page 10)



Book Notice



Forensic Speaker Identification

Author: Phil Rose

Phonetics Lab, The Australian National University

Expert opinion is being increasingly sought in the legal process as to whether two or more recordings of speech are from the same speaker. Forensic speaker identification (FSI) can be very effective, contributing to both conviction and elimination of suspects on the basis of their voice. Yet there is still a considerable lack of understanding on the part of law enforcement agencies, legal practitioners, and indeed phoneticians, speech scientists and linguists, as to what it involves; what constitutes appropriate methodology; what it can achieve; and what its limitations are. The aim of the book is to address these questions with the appropriate technical precision.

The book makes the points that forensic comparison of voice samples is extremely complex; that a Bayesian approach is the proper way of evaluating forensic speech samples; and that speech samples need to be compared both acoustically and auditorily, and from the point of view of their linguistic and non-linguistic features. The book also demonstrates that it can be done.

The book is organised around four

themes:

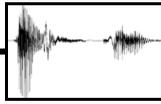
- (1) *The basic ideas in FSI*: what makes FSI difficult; forensic phonetic parameters; expressing the outcome - the Bayesian approach; characterising FSI.
- (2) *What speech sounds are like*: the vocal tract and the production and perception of speech sounds; phonemics; speech acoustics; speech perception.
- (3) *What a voice is*: a semiotic model for the information content in a voice.
- (4) *A demonstration of the method*: using Bayesian analysis on forensically realistic speech acoustics.

The summary addresses questions of what, in terms of data, method, and qualifications of the practitioner, constitutes requirements for a successful forensic speaker identification. It also asks what developments can be expected in the future.

Contents

1) Introduction. 2) Why voices are difficult to discriminate forensically. 3) Forensic phonetic parameters. 4) Expressing the outcome. 5) Characterising forensic speaker identification. 6) The human vocal tract and the production and

(Continued on page 10)



(Thesis abstract continued from page 8)

the overall likelihood ratio less than unity were interpreted as indicating a different-speaker pair, and values greater than unity were interpreted as indicating a same-speaker pair. Ninety non-contemporaneous same-speaker pairs, and 180 different-speaker pairs were tested.

Results showed that 5 out of the 180 different-speaker pairs were wrongly classified as same-speaker, and 9 out of the 90 same-speaker pairs were wrongly classified as different-speaker. This gives error rates of approximately 3% false alarms, and 10% missed hits. It was concluded that speakers can indeed be discriminated on the basis of their formant frequencies from non-contemporaneous natural speech with relatively few (i.e., 6) acoustic parameters. It was hypothesised that it might be possible to discriminate a greater number of speakers with more parameters.

Yuko Kinoshita

Research Assistant in
Forensic Phonetics
Linguistics
School of Language Studies (Arts)
The Australian National University.
Email: yuko.kinoshita@anu.edu.au
Phone: (02) 6125 8219

The 5 Column?

Have you got something to say about ASSTA? Let us hear your views. Please write to me, the Newsletter Editor:

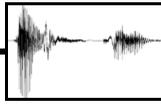
m.tyler@uws.edu.au

If you wish, also send me a brief biography. This is a good way to publicise your work and your views on Speech Science and Technology in Australia.

(Book Review continued from page 9)

description of speech sounds. 7) Phone-mics. 8) Speech acoustics. 9) Speech Perception. 10) What is a voice? 11) The likelihood ratio revisited: a demonstration of the method. 12) Summary and envoi.

For more details, see the *Taylor and Francis* web site: <http://www.tandf.co.uk/>



Speech Resources

by Roberto Togneri

We conclude the current series of speech resources from the World-Wide-Web with assorted links on speech and articulatory synthesis and auditory modelling.

Auditory Home Page

<http://www.auditory.org/>

Auditory Models

<http://www.auditorymodels.org/>

Articulatory Synthesis at Haskins Laboratories

<http://www.haskins.yale.edu/Haskins/MISC/ASY/ASY.html>

The Articulatory Database Registry

<http://www.cstr.ed.ac.uk/artic/>

Articulatory speech synthesis on the WWW

http://www.speech.kth.se/~olov/art_synth.html

The MBROLA Project

<http://tcts.fpms.ac.be/synthesis/mbrola.html>

Klatt Synthesiser Interface

<http://www.asel.udel.edu/speech/tutorials/synthesis/Klatt.html>

For a list of these and all WWW resources on speech science and technology known to me please stop by my Spoken Language Systems page:

<http://ciips.ee.uwa.edu.au/~roberto/research/speech.html>

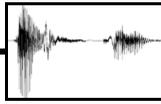
Or checkout the ASSTA WWW Links page:

http://www.assta.org/assta_sites.html

Have you come across any interesting URLs? Is there a classic book or journal in your area of research? Are you aware of useful software or speech data?

If so, please send your contribution to me at roberto@ee.uwa.edu.au

Contributions should contain a title, reference and brief description of the resource.



Conference Log

Conference log compiled by Marija Tabain

2002

Location: Manchester, UK

May 3 – 5

Information: <http://www.edgehill.ac.uk/acadepts/humarts/english/10mfm.htm>

*The Second International Conference
on Contrast in Phonology*

June 1

Location: Toronto, Canada

Information: <http://www.chass.utoronto.ca/~contrast/>

*LREC 2002 Workshop on Portability
Issues in Human Language
Technologies*

Location: Canary Islands, Spain

May 9 – 11

Information: <http://www.lrec-conf.org/lrec2002/>

*The Fifth Symposium on Natural
Language Processing 2002 + Oriental
COCOSDA*

June 3 - 7

*(Committee for the Co-ordination and
Standardization of Speech Databases
and Assessment) Workshop 2002*

*143rd Meeting of the Acoustical Society
of America*

Location: Hua Hin, Prachuapkirikhan,
Thailand

Location: Pittsburgh, Pennsylvania,
USA

Information: <http://kind.siit.tu.ac.th/snlp-o-cocosda2002/>

Information: <http://asa.aip.org/meetings.html>

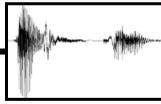
May 23 – 25

June 6 – 8

Tenth Manchester Phonology Meeting

*4th International Phonology Meeting of
the GDR 1954.*

Location: Grenoble, France.



Information: GDRphono@icp.inpg.fr

23 June - 6 July

NATO International Scientific Exchange Programmes: Dynamics of Speech Production and Perception.

Location: Il Ciocco, Italy

Information: <http://www.ebire.org/earlab/asi2002.html>

June 27 - 28

Conference on English Phonology

Location: Toulouse, France

Information: <http://www.edgehill.ac.uk/acadepts/humarts/english/toulouse.htm>

June 27 - 30

Laboratory Phonology 8 – “Varieties of Phonological Competence”

Location: New Haven, Connecticut, USA

Information: <http://www.ling.yale.edu/labphon8/>

June 28 – 29

Natural, Intelligent and Effective Interaction in Multimodal Dialogue

Systems

Location: Copenhagen, Denmark

Information: http://www.class-tech.org/events/NMI_workshop2.html

July 11

Workshop on Morphological and Phonological Learning

Location: Philadelphia, Pennsylvania.

Information: <http://morph ldc.upenn.edu/maxwell/MorphologyLearning.html>

August 24 – September 1

19th International Conference on Computational Linguistics- COLING2002

Location: Taipei, Taiwan.

Information: <http://www.coling2002.sinica.edu.tw/>

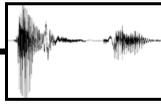
September 2 – 6

Linguistics and Phonetics 2002 (LP2002)

Location : Urayasu, Japan.

Information : midori34@meikai.ac.jp

September 11 – 13



IEEE Workshop on Speech Synthesis

Location: Santa Monica, CA, USA

Information: <http://www.research.att.com/conf/ttsworkshop/>

September 14 – 16

3rd Biennial ICVPB: International Conference on Voice Physiology and Biomechanics

Location: Denver, USA.

Information: <http://www.nwu.edu/csd/ICVPB/>

September 17 - 22

7th International Conference on Spoken Language Processing: ICSLP 2002

Location: Denver, Colorado, USA

Information: <http://www.icslp2002.org/>

September 14 - 16

3rd Biennial ICVPB: International Conference on Voice Physiology and Biomechanics

Location: Denver, Colorado, USA.

Information: <http://www.nwu.edu/csd/ICVPB/>

November 1 – 3

9th International Phonology Meeting: Structure and Melody

Location: Vienna, Austria

Information: <http://www.univie.ac.at/linguistics/conferences/phon02/>

November 30 - December 6

144th Meeting of the Acoustical Society of America, 3rd Iberoamerican Congress of Acoustics, 9th Mexican Congress on Acoustics

Location: Cancun, Mexico

Information: <http://asa.aip.org/cancun.html>

December 1 – 5

9th Australian International Conference on Speech Science and Technology – SST 2002

Location : Melbourne, Australia

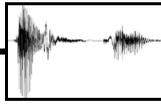
Information : <http://www.conferences.unimelb.edu.au/SST/>

2003

January 21 - 24

8th International Symposium on Social Communication.

(Continued on page 15)



5th Column

Multidisciplinary Musings

Author: *Michael Tyler, Macarthur Auditory Research Centre Sydney (MARCS), University of Western Sydney*

Our membership comes of a wide range of disciplines, such as Engineering, Computer Science, Linguistics and Psychology. I have attended SST conferences since 1996, and I am pleased to report that ASSTA members seem to embrace the multidisciplinary nature of Speech Science and Technology.

My experience at ICSLP2000 was somewhat different. There was a distinct air of segregation, which was highlighted when I met a British Engineer, who told me that he could not understand what psychologists and engineers had to do with Speech Science! As I have stated, this was not the case at SST, but I think there is still more that we can do to help each other out. For example, as an experimental psychologist, linguists can help me identify potential confounds in my stimuli, and provide a different perspective on my results. Engineers and computer scientists could suggest more efficient methods of analysis.

ASSTA members from all disciplines may be able to alert me to similar studies in their areas. For my part, I can give advice on experimental design, and answer questions on how humans process speech.

At this year's SST I look forward to hearing feedback about my results from any interested party!

What expertise can you share with other ASSTA members? Send your 5th Column article to me for the next issue!

(Continued from page 14)

Location: Santiago de Cuba, Cuba.

Information: leonel@lingapli.ciges.inf.cu

April 28 - May 2

145th Meeting of the Acoustical Society of America

Location: Nashville, Tennessee

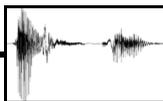
Information: <http://asa.aip.org/meetings.html>

August 3 - 9

15th International Conference of Phonetic Sciences

Location: Barcelona, Spain

Information: <http://shylock.uab.es/icphs/>



CALL FOR PAPERS

Submissions are invited for oral and poster presentations. Submissions should describe original contributions to spoken language, speech science and/or technology that will be of interest to an audience including scientists, engineers, linguists, psychologists, clinicians, and other professionals. The broad conference themes are:

- Human processing of spoken language.
- Machine processing of spoken language.
- Professional applications of speech science and technology.

Additional information may be found on the Themes page of the conference website: <http://www.conferences.unimelb.edu.au/SST/>.

Submissions will not be accepted for review later than 16th August 2002.

Authors have the choice of review on the basis of an extended abstract or the full paper. Reviews will be conducted anonymously by at least two members of the International Review Board. Authors will be notified of the review outcome no later than 13th September 2002. The final versions of accepted papers will be published as a CD-ROM and a printed abstract booklet. The booklet and the CD-ROM will indicate whether the paper was accepted on the basis of full paper or extended abstract review. We intend to invite the authors of the best full paper submissions (rated by the review board) to have their paper included in a special refereed journal issue which will be announced later. Final acceptance depends on:

- Favourable review by the conference peer review board.
- Receipt no later than 11th October 2002 of a final full paper and a final abstract in acceptable electronically readable format for inclusion in the conference proceedings.
- Receipt of payment for conference registration no later than 11th October 2002 from the submitting author.

Registration and submission forms, instructions for authors and templates will be available soon on the conference website, www.conferences.unimelb.edu.au/SST/.