

## EDITORIAL

**W**elcome to the second ASSTA Newsletter for 2010, and my first as editor. Thank you to Girija Chetty for allowing me the opportunity to take on this role, and for a great job over the last few years. Just to briefly introduce myself, I completed my PhD in sound processing for cochlear implants in 2009 at the Bionic Ear Institute and the University of Melbourne, and I now work for an online hearing aid company called Australia Hears. My main job is writing software that allows users to adjust their own hearing aids at home.

That's enough about me, let's get started with the newsletter! This issue includes a report on INTERSPEECH 2010 by an ASSTA travel award recipient, conference listings, and information about the upcoming Annual General Meeting.

In addition, our ASSTA foundation secretary and life member, Bruce Millar, has been tempted out of retirement to promote the role of automatic speech recognition to improve access to university lectures for students who have a language or hearing disability. Pages 8-9 explain the opportunity for ASSTA members with a wide range of interests and expertise to offer some level of assistance. Bruce would love to hear from any members to whom this may apply or who would simply like to know more.

Don't forget, this newsletter is the place to share information about events, conferences, lab profiles, and more. So please keep sending information and updates.

**Daniel Taft**

### ▶ Contents

President's Report	3
INTERSPEECH 2010	5
Liberated Learning	8
Conference Listing	10
AGM Announcement	11



# ASSTA INFORMATION

Australian Speech Science and Technology Association Inc.  
GPO Box 143, Canberra City ACT 2601

## ASSTA Executive

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

Secretary: David Grayden  
[grayden@unimelb.edu.au](mailto:grayden@unimelb.edu.au)

Treasurer: Shunichi Ishihara  
[Shunichi.Ishihara@anu.edu.au](mailto:Shunichi.Ishihara@anu.edu.au)

Member: Catherine Watson  
[c.watson@auckland.ac.nz](mailto:c.watson@auckland.ac.nz)

Member: Roberto Togneri  
[roberto@ee.uwa.edu.au](mailto:roberto@ee.uwa.edu.au)

## ASSTA Newsletter Team

Daniel Taft  
[daniel.taft@australiahears.com.au](mailto:daniel.taft@australiahears.com.au)

Roberto Togneri  
[roberto@ee.uwa.edu.au](mailto:roberto@ee.uwa.edu.au)

David Grayden  
[grayden@unimelb.edu.au](mailto:grayden@unimelb.edu.au)

## Contributions

Copy should be sent to the newsletter team via email as plain text or as an attachment in Rich Text Format (RTF) or Microsoft Word (DOC).

**Contributors:** Denis Burnham, Roberto Togneri, David Grayden, Erin Cvejic

## ASSTA Corporate Members

Appen Pty. Ltd.  
<http://www.appen.com.au/>

Cochlear Ltd  
<http://www.cochlear.com.au/>

CRC for Cochlear Implant and Hearing Aid Innovation  
<http://www.bionicear.org/crc>

Spectral Dynamics Pty. Ltd.  
<http://home.iprimus.com.au/elvoice/>

Visit ASSTA on the Web:  
<http://www.assta.org/>

# PRESIDENT'S REPORT

---



## ASSTA Aficionados

This year is an SST (Speech Science & Technology Conference) year, the first exclusive SST year since 2006 (in 2008 SST was rolled in with Interspeech). SST 2010 is from December 14-16 (Monday the 13th is a Tutorial Day), and is being hosted by La Trobe University at their City campus. 73 papers have been accepted and a good mix of these are by research students. As usual, the conference boasts low rego (max = \$440, \$275 for students) and three excellent keynote speakers: Professor Michael Robb from the Department of Communication Disorders, University of Canterbury; Professor Hugh McDermott from the Bionic Ear and the University of Melbourne; and

Professor Bob Ladd from the University of Edinburgh, who will also give a tutorial on designing materials for prosody experiments. In addition, there are now well-established special sessions in two areas that make us Australasian – PANZE (Phonetics and Phonology of Australian and New Zealand English) and OzPhon (Phonetics and Phonology of Australian Aboriginal languages). In addition, there is a special session on speech corpus data collection, which is topical for reasons set out below.

This year marks the start of two newly-funded ASSTA-supported Linkage grants, both of which incorporate speech corpus data collection. These are a Linkage Infrastructure, Equipment and Facilities (LIEF) grant, called the Big ASC project, in which auditory-visual speech data will be recorded in a variety of tasks at 12 or more locations around Australia<sup>1</sup>; and a Linkage Project in which speech corpus data focussed on forensic applications will be collected as part of a larger project. Both of these owe at least some of their success to the financial and moral support given by ASSTA (and many other institutions); it is good to see ASSTA serving a research support role. These two



corpora will be curated by ASSTA and will provide data for a range of speech science and technology applications for years to come.

The financial support of these projects marks a milestone in the now mature ASSTA Research Grants program, begun after ICSLP 1998. I urge you to think of ASSTA as a research partner and look beyond the first stage of ASSTA research funding (ASSTA Research Events) to the next step of securing external funds.

I think this is a vital and important role for ASSTA, and we would like to go on funding and promoting research in this way. In the last newsletter, I pointed out that our funds are not infinite and I posed a number of questions about future roles of ASSTA. The upcoming SST conference, both the AGM and corridor discussions, will provide a forum for this. See you there!

***Denis Burnham, November 2010.***

<sup>1</sup> It is rumoured that the hitherto closeted Black Box from the Big ASC project will make a guest appearance at SST.

# REPORT ON INTERSPEECH 2010

---

**Erin Cvejic**

MARCS Auditory Laboratories,  
University of Western Sydney

From the 27<sup>th</sup> to 30<sup>th</sup> September 2010, I attended Interspeech, the annual meeting of the International Speech Communication Association (ISCA), held in Makuhari, Chiba, Japan. Interspeech is the world's largest and most comprehensive conference on issues surrounding the science and technology of spoken language processing both in humans and machines, with a unique multidisciplinary approach drawing in a multitude of researchers from computer science, linguistics, psychology and engineering. The theme of the 11<sup>th</sup> Interspeech (which, as of the year 2000, has combined the International Conference on Speech and Language Processing, ICSLP, with EUROSPEECH) was "Spoken Language Processing for All: Ages, Health Conditions, Native Languages and Environments". This year, a total of 1324 papers were submitted, with 779 appearing in the conference program, and approximately 240 of these being given as oral presentations.

The conference proper began on Monday after a Sunday filled with a wide range of topics from "Speech and Language Technology for Linguists and Other Human Sciences" and "Mobile Voice Search", to "Medical Speech Processing". Based on the feedback from other attendees, these sessions were a worthwhile and knowledge bringing endeavour.

The conference was opened with welcome speeches by Keikichi Hirose and Satoshi Nakamura, as well as an entertaining and enlightening speech from Professor Hiroya Fujisaki, the inaugurator of ICSLP who in 1990 recognised the need for international collaboration among a myriad of scientific disciplines was essential to better understand spoken language processing and advance its technologies. Following the opening ceremony, Professor Steve Young, creator of the HTK toolkit for hidden Markov model-based speech recognition systems and the ISCA Medallist for 2010, gave the first of three keynote sessions in which he discussed the current state-of-the-art, illustrated with engaging corpus examples and outlined the potential future directions that statistically based spoken language technologies





## REPORT ON INTERSPEECH 2010

may take. Other keynotes throughout the conference were delivered by Professor Tohru Ifukube, who demonstrated ways that speech technology can benefit and enrich the lives of the physically challenged population, and Dr. Chiu-yu Tseng, who outlined the role of prosody in discourse processing for human listeners.

Interspeech ran four parallel oral sessions alongside four poster sessions at any one time, in addition to exhibitor stalls from local universities, Northern Digital Inc. (manufacturers of Optotrak), Google Research and Nuance. After a full first day, a welcome reception was held at the Hotel New Otani. Despite a shortage of food and an abundance of hungry academics, disappointment was largely avoided by drinks flowing late into the evening.

On the second day of the conference, I presented my work in an oral session entitled "Prosody: Analysis". While the other talks in the session dealt with semi-automatic analysis of pitch in speech corpora, my presentation outlined a way in which the visual correlates of prosody (e.g., eyebrow, rigid head, and hyperarticulated jaw movements) can be analysed using a

guided principal components analysis. The feedback received was positive, with interest from both Australian and international academics, with some helpful suggestions made for my future work. At the conclusion of the second day, a student-only banquet with an open bar was held, allowing for interaction in an informal and relaxed environment. A range of ISCA fellows were invited as special guests including Anne Cutler, Marc Swerts and Alan Black, who did the rounds inspiring PhD students from around the world.

Another day of keynotes, posters, and oral sessions covering diverse topics from Second Language Acquisition to ASR on the third day of the conference, which was completed by the highlight of the social program: the conference banquet. All attendees were issued with an Interspeech-branded wooden sake cup, which was used to enjoy all-you-can-handle sake refills from a self-service barrel. In addition to local cuisine, entertainment was provided by "Shuichiro and Daichi", a renowned shamisen band (a shamisen is a traditional Japanese three-stringed instrument played by plucking with a plectrum).



# REPORT ON INTERSPEECH 2010

Overall, the conference was of great benefit to attend, and provided me with the opportunity to see what other researchers both in my field and from other areas are getting up to, observing the state-of-the-art and getting feedback on my own work from respected academics. Next year, Interspeech will be held in Florence, Italy, from the 28<sup>th</sup> to 30<sup>th</sup> of August. I would like to thank ASSTA, ISCA and MARCS Auditory Laboratories for their individual contributions that have given me the opportunity to attend Interspeech this year, and I look forward to attending more Interspeech conferences in the future. See you in Florence!



# LIBERATED LEARNING CONSORTIUM



## LECTURE LANGUAGE MODEL DEVELOPMENT FOR AUSTRALIAN ENGLISH

The **Liberated Learning Consortium** is an international research group dedicated to improving access to lecture information for students with a disability such as deafness or other inefficient processing of language. Automatic Speech Recognition (ASR) is being researched as a tool to enable such access through captioning and transcription of lectures.

**ASR systems** require statistical models of both speech acoustics and language usage. The accuracy of such systems is enhanced when these models are based on an equivalent population of speakers and genre of language to the population and language domain of the eventual users of the technology.

**Disability services units** in Australian universities are required to provide access assistance for registered students and three of them are currently partners with the Liberated Learning Consortium (LLC). LLC would like to explore linkages in Australia with the speech science and technology community to accelerate the provision of an ASR system that could service needy students in Australian university lecture theatres.



## LIBERATED LEARNING CONSORTIUM

**IBM Research** has contributed a new, research grade, speech recognition engine (“Attila”) and implementation platform (Hosted Transcription Service) to the Consortium. The engine is configured for general transcription tasks, based on a U.S. English broadcast news corpus. Word error rates for an Australian English lecture have been measured at 35%, which is considerably higher than results for North American English lectures. This can be reduced by using statistical models derived from Australian lecture theatre conditions.

**European Media Labs** (EML, <http://www.eml-d.villa-bosch.de/>) have recently signed an agreement with LLC for the joint development of models including Australian English.

**Australian speech research community support is needed.** Many hundreds of hours of audio data from Australian lecturers are required. The target domain is the lecture environment (oral presentations that present spoken information about a

particular subject in a college or university setting). A start has been made with this task at Macquarie and ANU. In addition to carefully recorded audio data, verbatim transcripts and timing data are required to create the statistical models.

Opportunities for data collection and innovative expertise in orthographic transcription and segmentation to fulfill these requirements are sought.

For more information, please contact:

Dr J Bruce Millar,  
ANU College of Engineering and  
Computer Science  
[bruce.millar@anu.edu.au](mailto:bruce.millar@anu.edu.au)

Keith Bain,  
Saint Mary’s University,  
Liberated Learning Consortium  
[keith.bain@smu.ca](mailto:keith.bain@smu.ca)

Website:  
<http://www.liberatedlearning.com>



# CONFERENCE LISTING

## **ALTA 2010: Australasian Language Technology Workshop**

9 December 2010 – 10 December 2010  
Submission Deadline: 20 September 2010

Location: Melbourne, Australia

<http://alta.asn.au/events/alta2010>

## **ASRU'11: 2011 Automatic Speech Recognition and Understanding Workshop**

11 December 2011 - 15 December 2011  
Submission deadline: July 1, 2011

Location: Hawaii, USA

<http://www.asru2011.org/>

## **SLT 2010: IEEE Spoken Language Technology Workshop**

12 December, 2010 - 15 December 2010  
Submission deadline: Aug 2, 2010

Location: Berkeley, CA, USA

<http://www.slt2010.org>

## **SST2010: Thirteenth Australasian International Conference on Speech Science and Technology**

14 December 2010 – 16 December 2010  
Submission Deadline: 18 June 2010

Location: Melbourne, Australia

<http://www.assta.org/sst/2010/>

## **IWCS-2011: 9th International Conference on Computational Semantics**

12 January 2011 – 14 January 2011  
Submission Deadline: 30 September 2010

Location: Oxford, United Kingdom

<http://www.meaningfactory.com/iwcs2011/>

## **IADA 2011: 13th Conference of the International Association for Dialogue Analysis- Dialogue and Representation**

26 April 2011 – 30 April 2011

Submission Deadline: 1 September 2010

Location: Montréal, Québec, Canada

<http://www.dialogue-representation.org>

## **ICASSP 2011: IEEE International Conference on Acoustics, Speech and Signal Processing**

May 22, 2011 - May 27, 2011

Submission deadline: 20 October 2010

Location: Prague, Czech Republic

<http://www.icassp2011.com/>

## **INTERSPEECH 2011: 12th Annual Conference of the International Speech Communication Association**

28 August 2011—31 August 2011

Submission deadline: 31 March 2011

Location: Florence, Italy

<http://www.interspeech2011.org/>

## **WASPAA'11: 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics**

16 October 2011 - 19 October 2011

Submission deadline: TBA

Location: New Paltz, New York, USA

<http://www.waspaa.com/>

# AGM ANNOUNCEMENT

---

## **ASSTA Annual General Meeting - Notice and Agenda**

Australasian Speech Science and Technology Association Annual General Meeting  
La Trobe University City Campus, 15th December 2010

### Agenda

1. Attendance and apologies.
2. Confirmation of the minutes of the previous AGM held on 18th December 2009.
3. President's report
4. Treasurer's report
5. Report from the National Spoken Language Database Committee
6. Report from the Forensic Speech Science Committee
7. Election of officers.

The following nominations have been received for the five executive positions. As there are no contested office-holder positions, they are all elected automatically:

- President: Denis Burnham. Nominated by Shunichi Ishihara. Seconded by Roberto Togneri.
- Secretary: Catherine Watson. Nominated by Shunichi Ishihara. Seconded by Roberto Togneri.
- Treasurer: Shunichi Ishihara. Nominated by Roberto Togneri. Seconded by David Grayden.
- Executive Member: Roberto Togneri. Nominated by Shunichi Ishihara. Seconded by David Grayden.
- Executive Member: Marija Tabain. Nominated by David Grayden. Seconded by Catherine Watson

8. Any other business