EDITORIAL

Welcome to the second ASSTA Newsletter for 2013.

The President’s report from Denis highlights 3 valuable resources available to speech science researchers: ANDOSL, Austalk, and Alveo.

The “News” section of newsletter reports on the many successes of ASSTA members. It is great to share so please send to us any of your good news items. This issue see some new appointments and a retirement.

There is a report on the successful International Conference on Multimodal Interaction that was held in Sydney last December. ICMI received funding from ASSTA to assist students to attend the conference and the report contains a glowing review from one of the lucky students.

Two travel reports are included in this issue, both from the most recent Interspeech. Adele and Nicholas tells us of their experiences and it seems they enjoyed both the intellectual and social aspects.

Don’t forget, when you have news and information to publicise about speech science, this is the place to do it. So keep sending information and updates about your work and upcoming events.

Trent Lewis

Contents

ASSTA Information................................................................. 2
President’s Report............................................................. 3
News .................................................................................. 6
Flyers ................................................................................. 6
ASSTA Research Event Report................................................. 9
ASSTA Travel Reports........................................................... 10
Conference Listing............................................................... 13
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
Fax: (02) 9772 6736

Secretary: Marija Tabain
m.tabain@latrobe.edu.au

Member: Catherine Watson
c.watson@auckland.ac.nz

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

Member: Roberto Togneri
Roberto.Togneri@uwa.edu.au

ASSTA Newsletter Team

Trent Lewis
trent.lewis@flinders.edu.au

Member: Roberto Togneri
Roberto.Togneri@uwa.edu.au

Contributions

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

ASSTA Corporate Members

Appen Pty. Ltd.

Cochlear Ltd

The HEARing Cooperative Research Centre (CRC)
http://www.hearingcrc.org/

Spectral Dynamics Pty. Ltd.

Visit ASSTA on the Web:
http://www.assta.org/
The ASSTA-related corpora, ANDOSL and Austalk, and the virtual laboratory, HCSvLab (Alveo), are rich resources for research in speech science and the development of technological applications. See their websites and references given below for origins and further details. Herewith, a report on progress in and accessibility of these resources.

The Australian National Database of Spoken Language (ANDOSL)

andosl.rsise.anu.edu.au/andosl/

ANDOSL was an ASSTA initiative set up by Bruce Millar, Phillip Dermody, Jonathan Harrington and Julie Vonwiller funded by the Australian Research Council (ARC) and also from ANU and Macquarie.

ANDOSL data were collected in 1993-95 and consist of 200 read sentences, digit sets, /hVd/ words, and map task key-words and data, from 108 Australian speakers, 36 speakers of each of General, Broad, and Cultivated Australian English. Sentence recordings were also made of 9 speakers from each of the most populous migrant language groups – Italian, Greek, Cantonese, Serbian/Croatian, Vietnamese (Southern), Arabic (Lebanese), German, Polish, and Spanish (South American).

Access to ANDOSL has been cumbersome, but a recent change to the licensing agreement means that the ANDOSL corpus can be available over the network rather than having to mail out CDROMs, making distribution of the data much easier and removing the need to charge for duplication costs.

In the future, ANDOSL will be on the new Alveo (HCS vLab) virtual laboratory platform (see below) and work on importing that data will begin later this year. For now, there is a copy of the data hosted on a server at Macquarie University. Steve Cassidy, Steve.Cassidy@mq.edu.au, can provide access for download of the files.

Now that we can make the data more widely available, we will do some curation and make sure that we have as complete a version of ANDOSL as possible. We're aware that there have been a number of efforts to annotate parts of ANDOSL in various ways. We'd like to collect together as much of this as we can find and distribute it with the original data. If you have annotations on ANDOSL data that you can share, please contact Steve Cassidy.
President’s Report

The Big Australian Speech Corpus
(The Big ASC, AusTalk)
bigasc.edu.au

AusTalk is the product of the Big ASC project funded by the ARC, and contributions from 11 different universities and organisations (ASSTA included).

When complete, AusTalk will comprise 3 hours of audio-visual (AV) speech in various tasks (read sentences and words, digits, read and retold speech, spontaneous speech, yes/no variants, map task) from 1000 speakers of Australian English (AusE) from 16 sites across Australia. Data from the 200 speakers at the remaining 3 sites will be completed by the end of 2014. All data and metadata are available to project partners under the existing agreement. The audio data are available to other researchers and the video data will be made available as participants sign a new consent form that has been required by the lawyers. An exciting addition to AusTalk is the inclusion of 4 of the speakers who originally participated in the ANDOSL project (see above).

Annotation of Austalk commenced in late April 2012. The original goal to annotate all data will not be realised in this phase (due to funding shortfall), but we will provide protocols for further annotation by: a full set of manually created phonemic and orthographic transcriptions for a selected number of speakers; automatically time-aligned transcriptions for all the Read Speech data; and automatically generated orthographic transcriptions for a subset of the Spontaneous Speech data.

Follow-on projects using AusTalk have already begun including: data from further population groups at some locations; analysis of AusTalk data, e.g. video analysis for facial gestures and close phonetic analysis of the isolated word list data; using the Read Sentences as a set of stimuli to study listeners’ perceptual dimensions in characterising speakers’ vocal characteristics and speaking style. There will be a tutorial at Interspeech 2014 (Togneri, Bennamoun et al. 2014) where attendees will learn to use a 3D based AV corpus derived from AusTalk for AV speech/speaker recognition, and a presentation at the Language Resources & Evaluation Conference (LREC 2014; Estival, Cassidy et al. 2014).

The Human Communication Science Virtual Laboratory
(HCS vLab, Alveo)

Alveo is a virtual laboratory that is the product of the HCS vLab project funded by NeCTAR (National eResearch Collaboration Tools & Resources) along with contributions from 13 universities and 3 organisations (ASSTA included) most of which were also Big ASC project partners.

Alveo is a platform for easy access to language, speech music, text, and other communication science relevant corpora and the integrated use of a range of analysis tools on these corpora. It includes data storage, research computing services, and incorporates a web-based data discovery interface for searching and accessing data sets. Tool use is orchestrated by a workflow engine with both web and command line interfaces to allow use by technical and non-technical researchers. A paper on Alveo will be presented at LREC 2014 (Cassidy, Estival et al. 2014). The corpora and tools included in the first phase of the project are:
President’s Report

Corporea: PARADISEC (text, speech, video); AusTalk (AV); Australian National Corpus (text & speech); AVOZES (AV); UNSW music dataset (music excerpts from films); USYD room dataset (room impulse responses); MQ emotional prosody (speech); Colloquial Jakarta Indonesian corpus (audio & text); ClueWeb09 dataset. All these are now part of Alveo to some extent, with the following riders: AusTalk metadata is still being ingested and further data will be ingested as it becomes available; and ClueWeb data is indexed using INDRI and is accessible from the Research Cloud.

Tools: EOPAS (PARADISEC tool; text interlinear text, media analysis; NLTK (Natural Language Toolkit; text analytics; EMU (search, speech analysis, interactive labelling); AusNC Tools (KWIC, Concordance, Word Count, statistical analysis & summary on content subsets; Johnson-Charniak parser (full parse trees for text sentences); ParseEval (syllabic parse of consonant clusters); HTK (Hidden Markov Model Toolkit for speech recognition); DeMoLib (software for video analysis); PsySound3 (physical & psychoacoustical algorithms); ParGram (grammar for Indonesian); INDRI (information retrieval tool for large data sets). All have been integrated with 2 bonuses: the Johnson-Charniak Parser has been added to the NLTK so is now accessed similar to other text processing tools; and ParGram Indonesian grammar is available through the XLE interface, thus providing access to all the other XLE grammars.

Alveo Phase II (July 1, 2014 - June 30, 2016) will have 2 main foci. The first is Alveo uptake in and by the research communities which will be facilitated by regular training workshops and outreach sessions at relevant conferences. The first workshop is planned for mid-2014 in Sydney. Additionally, collaborations with other major projects are being sought; work in this regard has begun with PARADISEC and the new ARC Centre of Excellence, ‘The Dynamics of Language’. The second is adding selected resources and functionalities. Possible corpora include: the Liberated Learning Consortium (LLC) corpus of recorded lectures, an ASSTA-supported forensic database, and a sign language corpus; and tools include MAUS forced alignment with HTK and a new tool for music analysis.

Future Research in Australasian Speech Science and Technology

Alveo will house both ANDOSL and AusTalk and the tools by which to search and analyse these corpora; and also provide the infrastructure for storage, access and analysis of various other speech and other corpora. These developments are meeting the goals of Bruce Millar and his collaborators back in 1989, to provide bases for increased research in in Australian and now Australasian speech science and technology. (ANDOSL and AusTalk both involved only Australian English due to funding sources but it is hoped that New Zealand researchers can now build on this.) It is hoped that more researchers, both from Australia and internationally, will join Alveo, use its resources to generate research and knowledge and contribute new tools and new datasets. We eagerly await the emergence of projects that will take advantage of this new facility and which will contribute new results to our understanding of the science of human communication.

Denis Burnham,
ASSTA President,
April, 2014
Appointments

The University of New South Wales (UNSW), School of Electrical Engineering and Telecommunications is very pleased to welcome Dr Vidhyasaharan Sethu to its staff as a Lecturer. Dr Sethu's research interests include emotion recognition, speaker recognition, language identification and the application of machine learning in speech processing.

The School of Linguistics and Applied Language Studies at Victoria University (Wellington, NZ) is delighted to welcome Prof Miriam Meyerhoff to our School. Prof Meyerhoff is interested in language variation and change from a broad range of perspectives, with one major theme of her research being variation and change in creole languages. She uses a variety of approaches, including sociophonetics. Prof Meyerhoff has previously held academic positions at the University of Hawai at Manoa, Cornell University, University of Edinburgh and University of Auckland.

Retirement

Congratulations to Professor Andy Butcher on his recent retirement from Flinders University, where he spent 20 years as Foundation Professor of Communication Disorders.

Flyers

Workshop on the Role of Prosody in Language Learning: Stress, Tone and Intonation

Australian Hearing Hub, Macquarie University
8-9 Dec 2014

Sponsored by: ARC, Centre for Language Sciences (CLaS), Macquarie University

Much of the research on language acquisition has focused primarily on the segmental level of structure. Much less is known about how and when learners develop perceptual sensitivity to suprasegmental aspects of language, and how this development varies as a function of the language being learned. The use of stress, tone, and/or intonation varies widely across languages, with implications for word segmentation, learning the lexicon, syntax, semantics and discourse pragmatics, depending on the language. Using these prosodic cues to break into language, and putting them to practice in everyday speech, is critical to becoming a competent speaker of a language. Listeners also more easily and quickly comprehend the fluent use of prosodic structures, facilitating efficient and effective communication. However, monolinguals can be slow to exploit and produce prosodic structures, and L2 learning adults, or those with hearing loss, may never fully acquire these structures.

The goal of this workshop is to bring together researchers working on various aspects of stress, tone and intonation to discuss how these prosodic cues are exploited by learners, and put to use in language comprehension and production. The workshop will focus on both computational and empirical studies of both L1 and L2 language acquisition from a variety of languages and populations. It will include keynote addresses and
invited talks by experts in the fields of linguistics, cognitive science, computer science, and psychology. We also invite submissions for posters related to the workshop theme.

**Keynote Speakers**

Denis Burnham  
(MARCS Institute, UWS, Sydney)  
Anne Cutler  
(MARCS Institute, UWS, Sydney)  
Paul Smolensky  
(Cognitive Sciences, Johns Hopkins)

**Location**

The Workshop will be held in the Level 1 Theatre of the Australian Hearing Hub at Macquarie University, Sydney NSW, Australia.

**For further information:**

Email: wapst@mq.edu.au  
Web: [http://mq.edu.au/wapst](http://mq.edu.au/wapst)

---

**Important Dates**

*Poster submission deadline:* September 1, 2014  
300 word poster abstracts, including name and affiliation, should be emailed to wapst@mq.edu.au

*Poster Notification date:* October 1, 2014

*Registration deadline:* October 15, 2014  
Lunch and refreshments will be provided for registered attendees.

Please RSVP for catering purposes to wapst@mq.edu.au

**Organisation committee:**

Katherine Demuth, Mark Johnson, John Pate, Ivan Yuen, Nan Xu, Tamara Schembri & Gretel MacDonald
Call for Papers

The Australian Association of Speech Science and Technology is pleased to announce the call for papers for the 15th Australasian International Conference on Speech Science and Technology – SST 2014. The SST conference series is an international interdisciplinary conference designed to foster collaboration among speech scientists, engineers, psycholinguists, audiologists, linguists, speech/language pathologists and industrial partners.

We particularly invite submissions in the following areas: acoustic phonetics, audiology, clinical phonetics, speech and voice, forensic linguistics, acquisition, speech and signing pedagogy, phonetics and phonology of Australasian languages, sociophonetics, speech engineering and modelling, speech production and perception, speech prosody, speech science and technology applications, speech synthesis and recognition.

Dates: December 3-5 2014  
Deadline for submission: May 30, 2014  
Notification of acceptance: August 5, 2014

For further details:  
http://www.nzilbb.canterbury.ac.nz/SST.shtml

ASSTA New Researcher Award

ASSTA offers up to six New Researcher Awards for the best papers submitted for presentation at its SST conference. The award consists of a $750 grant, plus the SST conference dinner.

Applications close on the day the SST paper submissions are due.

Eligibility criteria:

(a) The applicant must be a Student member of ASSTA or an Ordinary member of ASSTA no more than two years beyond their highest degree and must be financial with their membership fees for at least 6 months at the time of application.

(b) The applicant must be conducting research in Australia or New Zealand.

For further details:  
http://www.assta.org/?q=assta-new-researcher-award
Final Report:  
ICMI 2013 Sydney  
Julien Epps, UNSW

The 15th ACM International Conference on Multimodal Interaction – ICMI 2013 – was held in Sydney, Australia during December 9-13th. Similarly to recent years, 2013 saw a very strong number of submissions: a record 133 long and short paper submissions, 11 demonstration papers, 5 multimodal grand challenge proposals and 3 workshop proposals. From the 133 long and short paper submissions, 49 were accepted for oral (27) or poster (22) presentation after a rigorous peer review process, bringing the acceptance rate to 37% overall and 20% for oral presentation.

The format of the conference comprised 3 Multimodal Grand Challenge Workshops, a Doctoral Consortium, a 3-day single-track main conference (including 3 keynote speakers, 7 oral sessions, 2 poster sessions and 2 demonstration sessions), and 3 Workshops.

Attendance at the conference was very strong: 167 researchers, research students and industry participants registered, which is the largest attendance of any ICMI conference outside of the US. Research students comprised 35% of the total registrants, which highlights the value of this conference to early career and emerging researchers. Attendees came from 34 different countries, the most strongly represented being the US, Japan and Australia.

Through the generous sponsorship of ASSTA, totalling A$2559, travel costs to attend ICMI were supported for three Australasian applicants, all research students: Jeewon Lee (RMIT), Shyam Rajagopalan (UCanberra) and Bin Liang (Charles Sturt University). Shyam and Bin also had their accommodation costs covered, due to their participation in the (peer-reviewed) Doctoral Consortium, which (via multiple funding sources) provided full funding for all participants, whether Australasian or international. All three expressed their appreciation for the support and for the development opportunities presented by the conference. Shyam commented on the “good amount of time given for every individual to present their research proposal and obtain *genuine* feedback from the [Doctoral Consortium] mentors, be it positive or negative. It certainly helps in refining the proposal, especially, if the student is in the initial stages of his/her PhD. The session during lunch from mentors on “Career Opportunities post PhD” was very informative and useful.”

From a speech science and technology perspective, there were a number of papers of interest across the various sessions and events, mostly related to either affective computing (recognition of emotion, expertise, deceit or depression from speech) or dialogue analysis/systems. The full proceedings may be accessed via the ACM Digital Library.

ASSTA was acknowledged through logos on the conference web site, on conference banners (at every site during conference), on conference bags and name badges, on background slides for all sessions, and on conference program; during the Chairs’ Welcome in conference proceedings; and verbally during the opening and closing speeches and during the conference banquet. The sponsorship certainly raised the profile of ASSTA among the international multimodal research community, with many participants asking more about ASSTA and its role and activities.
From the 25-29 August 2013 I attended Interspeech, the annual meeting of the International Speech Communication Association (ISCA), this year held in Lyon, France. Interspeech is the world's largest and most comprehensive conference on issues surrounding the science and technology of spoken language processing in both in humans and machines. It promotes an interdisciplinary approach covering all aspects of speech science and technology, ranging from fundamental contributions to practical applications, in a wide variety of areas. This year’s conference attracted over 1500 participants from all over the world. In addition it was also the occasion to celebrate the 25th Anniversary of ISCA.

Tutorials were conducted on the first day of the conference to provide information on advanced topics (recent advances) and also basic knowledge on a particular subject (crash course). I attended the tutorial "What speech researchers should know about video technology!" which provided information about recent advances in using large corpus video data (e.g. YouTube) in speech research.

Over the next four days Interspeech ran 6 parallel oral sessions alongside 3 poster sessions at any one time. This meant more than 50 oral and 30 poster sessions occurred during these four days. In addition the conference had several keynote speakers including Prof. Hynek Hermansky from Johns Hopkins University and this year’s ISCA medalist, who gave a presentation on “My adventures with speech”; Prof. Benjamin Munson from University of Minnesota who presented “On the interaction of social and linguistic factors in phonetic variation in typical and atypical speakers”; Prof. Anne-Lise Giraud from the University of Geneva who asked the question “Are cortical oscillations a useful ingredient of speech perception?”; and Maureen Clerc from INRIA who talked about “Verbal communication through brain computer interfaces”. There were also numerous Special Sessions and Show and Tell demo events, which allowed interaction between researchers, attendees and industry.

I had the opportunity to present my work in an oral session entitled "Speech acquisition and Development" on the second day of the conference proper. My presentation outlined the way in which the duration of infant vocalisations can be seen to have a non-linear developmental trend that is related to the use of non-modal voice qualities. This opportunity allowed me to meet and talk with fellow researchers about my studies. I received positive feedback and helpful suggestions for future directions.

Interspeech also held a “Student meet Experts” lunch where students in similar research areas had lunch together with two research experts. This was a valuable opportunity to meet other early career researchers who are working in similar areas and discuss our research together. The two experts on the table for “Speech production, perception, articulatory models and phonetics” were Prof. Catherine Best from MARCS and Assc. Prof. Slim Ouni from University of Nancy 2. Both were extremely friendly and interested in the work being conducted by the students around the table.
Interspeech also organized a comprehensive social program for conference attendees. There was a Welcome Reception, Students’ Reception and Conference Banquet that all provided good opportunities for informal mixing and socialising. The pre-Banquet event of “Slide Show Karaoke” was especially entertaining providing a huge number of laughs as both students and senior researchers took to the stage and presented on slides that they had no knowledge of and that had no internal coherency.

It was a great experience to attend this conference. The opportunity to hear recent advances, meet with other researchers and to present my work was invaluable. I wish to thank La Trobe University and ASSTA for their generous contributions in making attendance at this conference possible.

Nicholas Cummins
University of New South Wales

This year’s Interspeech was held in Lyon from the 25-29 August. On the Sunday I attended my tutorial “What speech researchers should know about video technology!” run by Koichi Shinoda and Florian Metze. Despite a moderate case of jet lag I enjoyed the tutorial and hope to be able to use some of the skills I learnt in practise soon. The plan for the rest of the day was to explore Lyon but my jet lag had different plans, I enjoyed a couple of hours of exploration around Vieux Lyon (the old town) before crashing out at the hotel for a sleep before meeting, speech research, friends for dinner.

The Conference proper started the on Monday with the welcome session and an excellent keynote address from Hynek Hermansky’s. I attended several interesting session including; Speech Analysis, a Deep Neural Network dominated ASR session, as well as the paralinguistic challenge sessions. For the welcome reception our French host built anticipation for our aperitifs nicely by first showing Julie Meitz’s speech research inspired video creation, http://vimeo.com/73783801, an explanation on the importance of aperitifs and had a band start playing, before lifting a curtain behind the main stage to reveal a selection of fine food and wine from Lyon. The least said about the resulting stampede of researchers desperate for the aperitifs the better!

My highlight on Tuesday was the paralinguistic analysis session; it was well attended session of really interesting talks, including my own! Other sessions I intended included; speech technology for speech disorders, speaker recognition, as
well as paralinguistic and speech analysis poster sessions. Again the quality of the talks and posters was outstanding. The days social event was the student reception, unfortunately due to a sore back I didn’t attend I was told the next day a good time was had by all those who attended, due in a large way to a (un)healthy amount of free wine, and possibly due to the amount of free drink there was a suspiciously large number of post docs present.

It was more of the same on the Wednesday with sessions on ASR, a speech science special session and poster session on paralinguistics taking up most of my day. Wednesday included the conference banquet and associated entertainment activities including Power point karaoke! For those who haven’t done Power point karaoke I suggest giving it ago, the session was the hardest I laughed, and been laughed at, ever at a conference! The banquet was fun, with large amounts of food and wine as well as entertaining trip through the music history of various styles of popular French and International music.

Attendance was well down for the first session on Thursday, a couple of the morning sessions had chairs who were either later or AWOL. The morning’s robust speaker recognition was the perfect place to ignore my wine induced headache and provided a couple of interesting papers as well as some laughs in the form of photos from the night before being included in one of the talks. A special event was held on after morning tea to celebrate 25 years of ESCA/ISCA which included talks by Joseph Mariani, Hiroya Fujisaki and Roger Moore on current and future directions of speech research and technology. My afternoon included some interesting posters sessions, the closing ceremony and for purely tourist reasons a walk related around the Parc de la Tête d’Or, the huge park the conference venue was located next to which included Lyon zoo.

Overall I enjoyed my time at Interspeech 2013, the venue was excellent and the program was well thought out minimising the need to be in two sessions at the one time. The use of both twitter for asking questions during keynotes as well as the program app were useful additions which I hope stay at future InterSpeeches. The inclusion of transport vouches in the welcome pack was great; Lyon is a larger city so this made travel that bit easier! I got lots of good feedback on my work, met lots of people, caught up with friends and learnt about lots of interesting speech research. I’m already looking forward to Interspeech 2014.
Conference Listing

SST 2014: 15th Australasian International Conference on Speech Science and Technology
December 3-5, 2014
Submission Deadline: May 30, 2014
Location: University of Canterbury, Christchurch
http://www.nzilbb.canterbury.ac.nz/SST.shtml

04 May 2014—09 May 2014
Submission deadline: Oct 27, 2013
Location: Florence, Italy
http://www.icassp2014.com/

19 April 2015 — 24 April 2015
Submission deadline: 5 October 2014
Location: South Brisbane, Australia
http://www.icassp2015.com/

September 14-18, 2014
Submission Deadline: 31 March 2014
Location: Singapore
http://interspeech2014.org

INTERSPEECH 2015: 16th Annual Conference of the International Speech Communication Association
September 6-10, 2015
Submission Deadline: 31 March 2014
Location: Dresden, Germany
http://interspeech2015.org

LabPhon14: The 14th Conference on Laboratory Phonology
July 25 — July 27, 2014
Submission deadline: Dec 1, 2013—Jan 31, 2014
Location: Tokyo, Japan
http://www.labphon.org/home/conferences

The International Conference on Multimodal Interaction, ICMI 2014
November 13th-15th, 2014
Submission Deadline: May 9th, 2014
Location: Bogazici University, Istanbul, Turkey
http://www.acm.org/icmi/2014/

The 18th International Congress of Phonetic Sciences (ICPhS2015)
Date: Aug 10 2015 — Aug 14, 2015
Submission Deadline: 1 Feb, 2015
Location: SECC Glasgow, Scotland
http://www.icphs2015.info/

TSD 2014 : Text, Speech and Dialogue
Date: Sep 8, 2014 - Sep 12, 2014
Submission Deadline: March 22, 2014
Location: Brno, Czech Republic
http://www.tsdconference.org/tsd2014/

LREC 2014 : The 9th edition of the Language Resources and Evaluation Conference
Date: May 28, 2014 - May 30, 2014
Location: Reykjavik (Iceland)
http://www.lrec-conf.org/lrec2014/

ACL 2014: The 52th Annual Meeting of the Association for Computational Linguistics
June 22, 2014 - June 27, 2014
Submission Deadline: Jan 10, 2014
Location: Baltimore, USA,
http://www.cs.jhu.edu/ACL2014/