



*Australian Speech Science
and Technology Association
(Incorporated)*

G.P.O. Box 143, Canberra

ASSTA Newsletter

January 2003
Volume 20, No. 4

Editorial

Welcome to the January 2003 issue of the ASSTA Newsletter.

You will notice in this issue that we have a new member of the editorial team - Lisa Stephenson will be looking after the conference log from now on (see page 5 for Lisa's profile). Welcome Lisa! Let me take this opportunity to thank Marija Tabain for her time as conference log sub-editor.

Towards the end of this issue, you will find two reviews of the SST conference, from the New Researcher Award recipients, Matt Flax and Marija Tabain. I have included Marija's review as part of the 5th Column because she has provided critical commentary on the SST conference series. In his president's report, Denis has encouraged all ASSTA members to voice their opinions on the association. The 5th



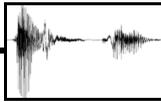
Column is one vehicle for this purpose.

Finally, Michael Wagner has kindly provided a profile of his laboratory in Canberra. We would like to highlight other labs in a similar way. If you are interested in contributing please contact me.

Michael Tyler

Contents

President's Report	3
ASSTA Initiatives	4
Sub-editor Profile	5
Lab Profile	6
PhD Abstract	8
Conference Log	9
Prix Benoît Photo	10
SST-2002 Review	11
5th Column	12



ASSTA Information

ASSTA Executive

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

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

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

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

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

ASSTA Newsletter Team

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

Sub-ed.(conferences): Lisa Stephenson
lisa@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 Lisa Stephenson, and thesis abstracts and book reviews to Johanna Barry.

The address for hard copy is:

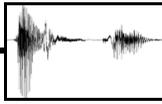
Michael Tyler
MARCS, University of Western Sydney
Bankstown Campus, Building 24
Locked Bag 1797, South Penrith D.C.
South Penrith, NSW, 1797.

Contributors: Matt Flax, David Grayden, Tom Johnstone, Maria Tabain and Michael Wagner. Thanks to Mel Gallagher and Karen Mattock for proofreading.

ASSTA recognises the assistance of the *MARCS Auditory Laboratories* 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



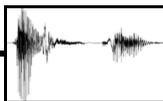
This newsletter contains two reviews of SST-2002, held in Melbourne, December 2 to 6, 2002, both by ASSTA New Researcher Awardees. One sings the praises of SST, especially its diversity, and the size - small enough to get around and meet people; the other, presented in the 5th Column, is more critical, suggesting ways in which SST could be improved. There are two interconnected themes I would like to pick up: the centrality of SST to ASSTA, and the way forward for ASSTA via critical input from members.

The Australian International Conference on Speech Science and Technology is the premier meeting on speech science and technology in Australia. SST conferences are student friendly and attract a fair sprinkling of participants from outside Australia. They continue to be one of the defining char-

acteristics of ASSTA. While in no way minimising this role, note that SST is concerned with the *products of research*, the research report or paper. True, at SST we can talk to colleagues about research future directions, and meet new potential collaborators. However, ASSTA must also be directly involved in *shaping* and facilitating speech research in Australia, and use its unique position as a bridge between science and technology to engender links between academia and industry. We must be involved in *promoting research*.

The ASSTA research initiatives have been designed with this in mind. In particular, there is ASSTA funding (going begging) in the form of Research Events. We have had one (successful) application for such funding so far, by a consortium of relatively established, and much newer ASSTA members. ASSTA Members can apply to ASSTA for funds to facilitate their research, and shape the future direction of speech research in Australia. Pay your \$40 annual fee and get much more than a few lifetimes of fees in return in research funds, and much much more in terms of research output, involvement of students, international collaboration, or whatever it is that your research event is directed at.

(Continued on page 4)



(Continued from page 3)

ASSTA is ready to move forward. We need to establish ourselves as the peak body for speech science and technology in Australia, and this involves all stages of the research process – attracting students, facilitating research ideas and collaborations, displaying our research at local and international conferences; interchange between academia and industry; and promoting our research both within speech-related disciplines, and to the wider community.

And so, to the point about critical input. Next time you find yourself thinking, “ASSTA is not relevant to my research”, stop and either write an email to or call one of the executive members (see page 2 for this information), pen a note for the 5th column in the ASSTA Newsletter, or write an application for money to fund a research event that *is* relevant - discussion between potential collaborators, a meeting with a potential industry partner, a small seminar, whatever meets your speech research interests. ASSTA members do research in speech science and technology. ASSTA is not separate from its membership; it promotes research in speech science and technology. Help us to hone the ways in which ASSTA does this.

ASSTA Initiatives

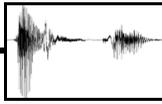
ASSTA manages a number of initiatives that aim to provide support to members in their research in Speech Science and Technology. The initiatives are:

- **The New Researcher Award Scheme** for the most promising projects presented to its SST conference.
- **PhD Study Awards** for student members beginning a PhD.
- **ASSTA PhD of the Year Award** for the best PhD completed each year.
- **Research Event Awards** for research events organised by ASSTA members.
- **ASSTA Travel Awards**, for travel to conferences.
- The **SST Conference Series** to provide a national forum for the dissemination of our science and technology.
- The **National Lecture Tour Programme** to enable prominent experts to lecture in provincial areas.

ROUND 1 for 2003 has commenced: applications in the first round for PhD Study Awards and Research Event Awards are due on 30 May 2003. If you are a PhD student or supervisor, or you wish to apply for funding for a research event, don't miss this opportunity!

Details of these initiatives are available at the ASSTA website: <http://www.assta.org/initiatives/>

David Grayden, Executive Member



New Sub-editor Profile



Lisa Stephenson

Macquarie Centre for
Cognitive Science

I came to Macquarie University in 1996, enrolling in a 4 year Bachelor of Psychology (Hons) degree. In 2000, I undertook the Psychology honours program under the supervision of Veronika Coltheart, investigating the effect of various psycholinguistic variables and presentation rate on comprehension.

In 2001, I enrolled in a PhD in Psychology at Macquarie. I am investigating aspects of spoken language involved in speech comprehension with Max Coltheart, Marija Tabain and Jonathan Harrington. I also work as a casual research assistant for Associate Professor Veronika Coltheart.

Research Details

Thesis Title : Assimilation in English and Japanese.

The general area of my thesis is the human ability to communicate. More specifically, I am interested in how speech is produced and processed. Firstly, I started looking at the process of

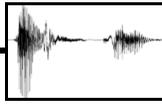
assimilation – changes in the production of a particular sound as a function of the surrounding phonetic context. I have investigated how assimilation may differ between English and Japanese. By looking at the patterns of assimilation across different languages, we can gain insight into how cognitive representations of language interact with the process of speech production. Analyses were based on acoustic and electropalatographic (EPG) data (electropalatography is a method of measuring contact between the tongue and the roof of the mouth and a unique palate is required for each speaker).

To investigate further the relationship between cognitive representations and speech production, my next study explores how lexical frequency of compound words may influence their production. Past work has shown effects of lexical frequency on speech perception and so this study will also investigate whether effects are found in articulatory production using EPG data.

Conference log

I will now be looking after the conference log, so if you come across a new conference that you think others will be interested in then email me the details to lisa@maccs.mq.edu.au.

Thanks!!



Lab Profile



*The Human-Computer
Communication Lab*

University of Canberra

The Human-Computer Communication (HCC) Lab at the University of Canberra is part of the University's School of Information Sciences and Engineering (ISE) within the Division of Business, Law and Information Sciences. HCC Lab has 20 members, most of whom are concerned with speech technology applications, their enabling technology and underlying theory.

Currently, the Lab has active research projects in the six application areas shown in the diagram on the next page. In the application areas of dictation systems and telephone query systems, research projects have recently combined the expertise of the speech recognition group (Wagner, Tran, Davies) and the human-computer interaction group (Collings, Walker, Webb, Kraal, Lynch). In the application areas of telephone and Internet authentication systems, current projects include the development of fuzzy and evolutionary techniques and their incorporation in conventional acoustic models, and the fusion of facial information and voice information for a powerful new face-voice authentication paradigm (Tran, Wagner, Le, Huang).

At the level of underlying theory, current

work in progress includes research on task models (Davies, Webb), language models (Davies, Wagner), hidden Markov models (Tran), fuzzy methods (Tran, Le), artificial neural networks (Clark, Le), face models (Huang, Wagner), cryptography (Sharma), and digital signal processing (Whichello, Stone).

Research projects in several of the application areas are undertaken by postgraduate research students of the School of ISE at PhD and Master levels, as well as by Bachelor of Computer Engineering and Bachelor of Software Engineering students who undertake year-long final-year group projects in one of these areas.

The HCC Lab is equipped with a Sun Sparcstation 2, an Aculab Prosody telephone interface, and a number of state-of-the-art PCs with audio and video input facilities. The Lab has various software development platforms such as the Hidden Markov Model Toolkit and MS Visual Studio, and it has access to a large number of spoken language data corpora through the Linguistic Data Consortium.

The HCC Lab researchers in the speech and language areas are:

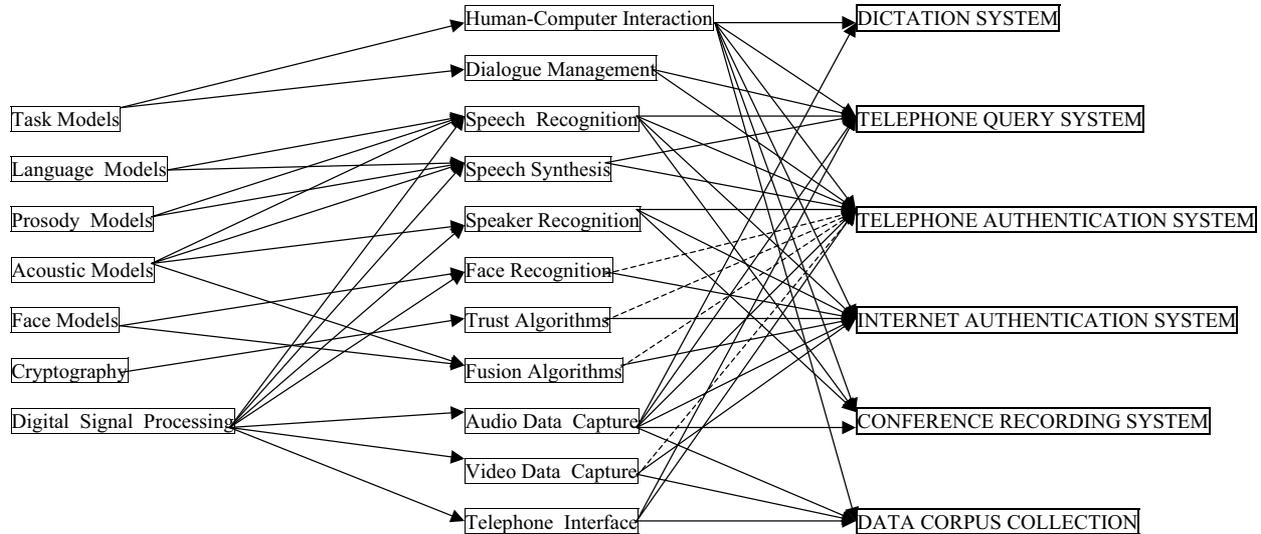
Prof. Michael Wagner, Dr David Clark, Penny Collings, Dr Dave Davies, Dr Xu Huang, Ben Kraal, Dr Kim Le, Neil Lynch, A/Prof. Van Le, Dr Dharmendra Sharma, A/Prof. Brian Stone, Dr Dat Tran, Dr David Walker, Dr Terry Webb (DSTO), and Dr Adrian Whichello.

SPEECH TECHNOLOGY APPLICATIONS AND ENABLING TECHNOLOGIES

Underlying Theory

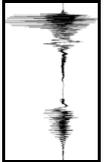
Enabling Technology

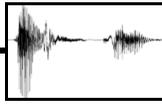
Application



Selected References

- Dat Tran and Michael Wagner, Generalised Fuzzy Hidden Markov Models for Speech Recognition, Lecture Notes in Comp Sci: Advances in Soft Computing - AFSS 2002, N.R. Pal, M. Sugeno (Eds.), pp. 345-351, 2002.
- Dat Tran and Michael Wagner, A Fuzzy Approach to Speaker Verification, Int J Pattern Recognition and Artificial Intelligence (IJPRAI), vol. 16, no. 7, pp. 913-925, 2002.
- Dat Tran and Michael Wagner, A Proposed Likelihood Transformation for Speaker Verification", Proc the Int Conf on Acoustics, Speech & Signal Proc (ICASSP), vol. 2, pp. 1069-1072, Istanbul, Turkey, 2000.
- Penny Collings, David Walker and Michael Wagner, Developing mental models and new work practices: An evaluation of a state-of-the-art commercial speech recognition system, Human Factors Conference HF2002, Melbourne, 2002.
- Ben Kraal, Michael Wagner and Penny Collings, Improving the User Interface of Dictation Software, Proc. 9th Austr Int Conf on Speech Sci & Tech, Melbourne, pp 22-27, 2002.





Ph.D. Thesis Abstract



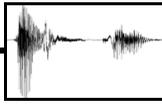
Ph.D. Thesis: *The effect of emotion on voice production and speech acoustics*

Author: Iain T. Johnstone, Department of Psychology, University of Western Australia

The study of emotional expression in the voice has typically relied on acted portrayals of emotions, with the majority of studies focusing on the perception of emotion in such portrayals. The acoustic characteristics of natural, often involuntary encoding of emotion in the voice, and the mechanisms responsible for such vocal modulation, have received little attention from researchers. The small number of studies on natural or induced emotional speech have failed to identify acoustic patterns specific to different emotions. Instead, most acoustic changes measured have been explainable as resulting from the level of physiological arousal characteristic of different emotions. Thus measurements of the acoustic properties of angry, happy and fearful speech have been similar, corresponding to their similar elevated arousal levels. An opposing view, the most elaborate description of which was given by Scherer (1986), is that emotions affect the acoustic characteristics of speech along a number of dimensions, not only arousal. The lack of empirical data supporting such a theory has been blamed on the lack of sophistication of acoustic

analyses in the little research that has been done. By inducing real emotional states in the laboratory, using a variety of computer administered induction methods, this thesis aimed to test the two opposing accounts of how emotion affects the voice. The induction methods were designed to manipulate some of the principal dimensions along which, according to multidimensional theories, emotional speech is expected to vary. A set of acoustic parameters selected to capture temporal, fundamental frequency (F0), intensity and spectral vocal characteristics of the voice was extracted from speech recordings. In addition, electroglottal and physiological measurements were made in parallel with speech recordings, in an effort to determine the mechanisms underlying the measured acoustic changes.

The results indicate that a single arousal dimension cannot adequately describe a range of emotional vocal changes, and lend weight to a theory of multidimensional emotional response patterning as suggested by Scherer and others. The correlations between physiological and acoustic measures, although small, indicate that variations in sympathetic autonomic arousal do correspond to changes to F0 level and vocal fold dynamics as indicated by electroglottography. Changes to spectral properties, speech fluency, and F0 dynamics, however, can not be fully explained in terms of sympathetic arousal, and are probably related as well to cognitive processes involved in speech planning.



Conference Log

Conference log compiled by Lisa Stephenson

2003

April 7 – 9

March 27 - 29

*Ecole thematique de phonologie et de
phonetique*

Location: Ile-de-Porquerolles, France

*International Colloquium on Prosodic
Interfaces*

Information: [http://www.lpl.
univ-aix.fr/~etpp03/](http://www.lpl.univ-aix.fr/~etpp03/)

Location: Nantes, France

Information: [ip2003@humana.
univ-nantes.fr](mailto:ip2003@humana.univ-nantes.fr)

April 10 – 12

April 6 – 10

*39th Annual Meeting of Chicago
Linguistic Society*

Location: University of Chicago, USA

*IEEE International Conference on
Acoustics, Speech and Signal
Processing*

Information: [http://
humanities.uchicago.edu/
orgs/cls/](http://humanities.uchicago.edu/orgs/cls/)

Location: Hong Kong

Information: [http://www.eie.
polyu.edu.hk/~icassp03/](http://www.eie.polyu.edu.hk/~icassp03/)

April 12

April 7 - 9

Boundaries in Intonational Phonology

Location: Lund, Sweden

*Eighth Western Pacific Acoustics
Conference*

Information: [http://glow.kub.
nl/Lund/phon_cal.htm](http://glow.kub.nl/Lund/phon_cal.htm)

Location: Melbourne, Australia

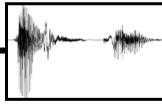
April 18 – 20

Information: [http://www.
wespac8.com/](http://www.wespac8.com/)

*Phonetics Today: 4th International
Conference*

Location: Moscow, Russia

Information:
phoneticstoday@yandex.ru



April 28 - May 2

*145th Meeting of the Acoustical Society
of America*

Location: Nashville, Tennessee

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

June 9 - 11

*4th International Conference on Audio
and Video Based Biometric Person
Authentication*

Location: Guildford, United Kingdom

Information: <http://avbpa2003.ee.surrey.ac.uk/>

July 1 - 4

2003 Child Phonology Conference

Location: Vancouver, Canada.

Information:

stemberg@interchange.ubc.ca

July 5 - 6

*4th SIGdial Workshop on Discourse and
Dialogue*

Location: Sapporo, Japan

Information: <http://www.speech.cs.cmu.edu/sigdial2003/>

July 7 - 11

From representations to constraints

Location: Toulouse, France

Information: phon2003@univ-tlse2.fr

August 3

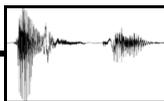
*Intonation in language varieties: AM
approaches*

Location: Barcelona, Spain

Information: <http://www.vuw.ac.nz/lals/icphs/>

Johanna Barry speaking at ICSLP2002 after having received her Prix Christian Benoit. See Volume 20, No. 2 for more details of the award.





SST2002 Review



Author: Matt Flax

University of New South Wales

Without a doubt one of the most interesting facets of the SST conferences is its cross-disciplinary nature: Linguistics, Cognitive Neuroscience, Speech Processing, Signal Processing, Databases, software interfaces and phonetics (to name but a few). People from around the world gathered in December 2002, some more ready than others to brave Melbourne's four seasons in one day.

The character of SST2002 reflected its physical surrounds. Australia, famous for its quirky continuity, kept the extra-curricular activities thick and fast. Melbourne, famous for its laid-back attitude, enveloped the conference with a relaxed aura.

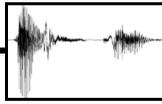
Some keynote speakers dealt with the mechanisms of hearing, understanding and articulating speech. Others dealt with the evolution of language (English in this case) and technology which aids the blind. It was possible to walk from session to session and gather concepts from engineering and science, ranging from psychology to people-made systems. Highlights included research delving into triphone (consonant-vowel-consonant) Thai, concepts of isolating meaning simply from prosody, software for research tools, models of cognitive processes, and forensics.

SST conferences are currently small (<100 publications). This allows people to meet from all areas of research. It also allows young researchers to mix with wiser researchers, who with a few words can influence project directions and generate insight. The cross pollination of ideas give people the opportunity to think beyond immediate directions in their research field and affirm broader 'nature of the world' beliefs.

(Continued from page 12)

conference. Personally I am always tired and grumpy at the end of a long day listening to talks, and I find that I have less enthusiasm to tackle a poster session than I would earlier on in the day.

I don't have any strong views on whether SST should be combined with other Australian conferences related to speech. If such a combination would mean that SST conferences take place more often than every 2 years, I would be keen for this to happen. I do realize that combining conferences would be a logistical nightmare, but I would welcome the opportunity to attend keynotes and selected papers that would normally be presented at the NLP, acoustics or psycholinguistics conferences. My main focus would still be on the speech papers, but I always enjoy finding out what is happening in other related research areas.



5th Column

Thoughts on the SST conference

Author: *Marija Tabain, Macquarie University.*

The people at ASSTA committee were very kind to offer me a New Research Award at the recent SST meeting in Melbourne. I was particularly sensitive to this kindness since I'd won the award once before, in Adelaide in 1996. Since then, the award guidelines have changed, so that Early Career Researchers (ECRs), and not just PhD students, can apply. Since ECRs are my favourite special interest group, I am very grateful to ASSTA for offering me the award again, and I hope that ECRs will continue to benefit from the fiscal generosity of ASSTA.

In order to express my gratitude more publicly, Denis and Michael asked me to contribute this note to the ASSTA newsletter, to reflect on the conference and perhaps on the role of ASSTA for Australian speech research. Ever eager to put forward any ill-informed opinions I may have, I herewith express my views on the SST conferences.

I believe the following issues need to be considered:

(1) Is Tutorial Day really worth it? The only reason I attended the Tutorial Day this time around was because I was presenting. I noticed that the number of attendees was quite small compared to the main part of the conference. I imagine that organizing presenters is quite tedious, and hiring the venue for an extra day is quite a cost.

(2) Do we really need full papers? We all know that writing a 6-page paper is quite a bit of work for the author(s). It also adds quite a load to the conference organizing task. Perhaps a model like that of Language and Speech, where only abstracts are submitted (which are then published in an Australian scientific journal), is the way to go. I realise that this deprives students of an opportunity to write a full paper, and that some more senior researchers welcome the opportunity to write a full conference paper (either for the peer-review status, or because they wouldn't have time to write up the results as a full journal paper), but I personally would be happy not to have to write a full paper for SST.

(3) I think poster sessions should be rethought. There were too many posters to view during the one session. If we reduced the overall number of papers by accepting only one first-author paper per attendee, and perhaps considered having more than one poster session, the outcome may be more satisfactory. In addition, I think poster sessions should not be tacked on at the end of the day, as an afterthought to the main part of the

(Continued on page 11)