



update

what's happening in speech
and language technology in Australasia

CONTENTS

- Welcome
- Focus:
LangTech2003 Report
- The Australasian Language
Technology Workshop and
Summer School
- Heard on the Street
- Upcoming Events

Welcome to the fifth issue of LT Update!

LT Update is a free publication from the Centre for Language Technology, produced with the generous support of CSIRO. The Update is a twice-yearly hard and soft copy publication, backed up by timely email alerts, that aims to keep you abreast of developments in the speech and language technologies in Australia and New Zealand.

If you're not yet a subscriber, sign up at <http://www.clt.mq.edu.au/LTUpdate>. If you are a subscriber and you want to change your subscription details, visit the site and key in the six-character passcode printed on the top of your mailing label.

In this issue, we provide a report on LangTech 2003, held in Paris last November. Plus, brief reports on the Australasian Language Technology Association's workshop and summer school, and news and events from the language technology world.

You can find out more about the topics in this issue via the links at our website: visit <http://www.clt.mq.edu.au/LTUpdate>.

What's your view?

If you have comments on LT Update, or ideas on things you'd like to see us cover, just mail ltupdate@ics.mq.edu.au.

LangTech 2003

In each issue of LT Update, we focus on a theme or issue of importance in speech and language technologies. In this issue, Andrew Joscelyne reports on LangTech 2003, the main industry event for speech and language technologies in Europe.

LangTech is Europe's very own industry-focused language and speech technology conference cum networking event. The second such get-together was held in Paris on November 25 and 26 2003, organised by ELRA-ELDA (www.elra.info), and attracted some 350 attendees, a majority of them from industrial/commercial language technology organisations. The exhibition areas featured some 21 stands, ranging from Multilingual Computing & Technology magazine (www.multilingual.com) to a large contingent from France Telecom's R&D division.

LangTech evolved as an event in 2002 out of the now-ended EUROMAP project (funded by the European Commission; www.itri.bton.ac.uk/projects/euomap) dedicated to raising awareness about language technology opportunities. It is supervised by an informal organising committee including senior officers from the Copenhagen Center for Language Technology, the DFKI (Germany), and ELRA (France).

The talks programme covered new and emerging language/speech applications in three fields: speech, knowledge management and multilingual communication. The speech topics ranged from mainstream speech-driven call centre type applications to a coming-to-market speech translation system. Knowledge talks focused largely on existing and emerging search and classification applications (patents for example), as well as an interesting proposal from Hans Uzsköreit on link-driven associative memory as a stage on the way to the semantic web where language technology might play a role. Multilingual applications

(always a slightly difficult category since multilinguality is a horizontal rather than vertical property of language systems) featured practical examples of translation and multilingual search solutions in a number of well-explored domains.

The keynotes were given by Stuart Patterson, CEO of ScanSoft, and Sanda Harabagiu, director of Language Computer Corporation in Texas. Patterson gave a broad overview of where commercial speech technology was going. He showed how ScanSoft was starting to move beyond generic speech technology development to focus on specific applications areas such as Directory Assistance language models. Harabagiu gave a detailed presentation of a new Q&A system under development in her company, mainly for government intelligence clients, showing how semantic knowledge was integrated into her solution.

LangTech also ran an elevator pitch session for 12 start-ups or small companies from 10 countries. Each 'competitor' had five minutes to present their company to a jury of venture capitalists and language tech professionals. The LangTech 2003 Prize for the three best presentations went to Lingua & Machina (France) which is developing next generation translation memory software, Morphologic (Hungary), well-established in developing language technology for Eastern and Central European languages, and CIMOS (France) which is seeking funding for the development of a Multilingual Universal Semantic Analyzer engine. Among other competitors was an innovative language learning project from Germany offering a platform for integrating speech technology apps (though only one other presenter (Canoo from Switzerland) talked about language-learning technology applications), and a new speech translation application targeting healthcare market needs in the multicultural USA.

continued overleaf ...

ALTA Workshop and Summer School

In December 2003, the Australasian Language Technology Association (ALTA; www.alta.asn.au) hosted a very successful week-long summer school and workshop programme in Melbourne. Introductory courses were offered on Practical NLP using Python, Speech processing, Dialogue systems, and Information extraction and question answering; advanced courses covered Machine translation, Validation and evaluation in NLP and IR, Statistical parsing, and SVMs and kernel methods in NLP. All the materials from the courses are available from the web site at www.conferences.unimelb.edu.au/alta/. The one-day workshop in the middle of the summer school showcased language technology research activity around Australia and beyond via 18 technical papers.

Heard on the Streets

In a pilot roll-out that started in August last year, Qantas Frequent Flyers across Australia can take advantage of a speech-enabled customer self-service application called Qantas Voice. This system allows callers to make Australian domestic award bookings, check their points balance, calculate how many points are needed for reward travel, and check award flight availability. The app was developed by Avaya and ScanSoft.

ScanSoft is also behind the recently deployed Centrelink voice recognition system which allows customers to report earnings income. Every fortnight, 200,000 customers have to provide this information, generally via interaction with a live operator. Within five weeks of deployment, 30,000 users had opted in to use the speech system, compared to 15,000 for the internet channel, and Centrelink expects up to 100,000 customers to move to using the speech channel.

Who gets LT Update?

LT Update is a product of Macquarie University's unique teaching program in the human language technologies. This program, funded under the Federal Government's prestigious Science Lectureships Initiative, is the only teaching program in Australia that focuses on delivering a rich education in the twin areas of spoken language processing and natural language processing, widely viewed as critical technologies in the next few years. LT Update is provided as a service for alumni from this program, so it provides both a community for those with similar interests, and at the same time a very focussed channel to a group of people with particular skills. Thanks to CSIRO's generous support, subscriptions are currently free: visit <http://www.clt.mq.edu.au/LTUpdate> to register. You can also access this newsletter electronically via site and you'll also find there web links to all the items mentioned in this issue as well as pointers to further resources.

Editor:

Robert Dale (rdale@language-technology.com)

LangTech 2003 *continued ...*

Panel sessions broached (but did not unfortunately have time to discuss) the difficulties of marketing language and speech technology, the quantitative state of the market in Europe, the future of funding in Europe (European Commission and French and German government proposals, each very different in scope and purpose). A final round table in honour of the late Antonio Zampolli (a major influence in EU language technology who died in July 2003) was dedicated to the future: better speech, more multimodality, ... but no one put a date to milestones on the road map.

LangTech 2003 was held at a difficult time for suppliers (the new round of European IT programme funding hadn't started,

technology spend is down, and many small language technology companies are fighting for their lives). It was generally agreed that language technology customers want integrated solutions, not technology. All of which requires a huge marketing effort to make these technologies more familiar to potential customers and users, and compelling arguments that language and speech applications are worth investing in.

Check out the LangTech 2003 website at www.lang-tech.org for downloads of all programme and elevator pitch slide shows.

Andrew Joscelyne is a Paris-based writer on language technology and member of the LangTech Organising Committee.

Upcoming Events

National

- **Second Annual Voice World Conference: 9-11 March 2004**, Convention Centre and Exhibition Centre, Sydney. <http://www.customercontactworld.com>.
- **Fourth Annual Conference for Standards and Process in Publishing (Open Publish 2004): 28-30 July 2003**. All Seasons Premier Menzies Hotel, Sydney. <http://www.openpublish.com.au>.

International

- **Sixth Annual Telephony Voice User Interface Conference: Putting Speech to Work in Business: 4-6 February 2004**, Maimi Hilton, Florida. <http://www.tmaa.com/conference>.
- **IJC-NLP04, the First International Joint Conference on Natural Language Processing: 22-24 March 2004**, Hainan Island, China. <http://www.cipsc.org.cn/IJCNLP-04/>.
- **Workshop on Spoken Language Understanding for Conversational Systems at HLT/NAACL 2004: 7 May, 2004**, Boston, Massachusetts. <http://www.research.att.com/~dtur/NAACL04-Workshop/>.
- **4th International Conference on Language Resources and Evaluation (LREC 2004): 26-27 May 2004**, Lisbon, Portugal. <http://www.lrec-conf.org/lrec2004/index.php>.
- **5th ISCA Speech Synthesis Research Workshop: June 14-16th 2004**, Carnegie Mellon University, Pittsburgh, USA. <http://www.ssw5.org>.
- **42nd Annual Meeting of the Association for Computational Linguistics (ACL2004): 22-24 July 2004**, Barcelona, Spain. <http://www.acl2004.org>.
- **8th Pacific Rim International Conference on AI (PRICAL): 9-13 August 2004**, Auckland, New Zealand.
- **The 20th International Conference on Computational Linguistics (COLING 2004): 23-27 August 2004**, University of Geneva, Switzerland. <http://www.issco.unige.ch/coling2004>.