

## Tiny Voices

Most of the TTS systems out there are intended for running on 'real computers', but Cepstral of Pittsburgh are making sure that the small end of the computing market isn't left behind: check out their website at <http://www.cepstral.com> for demos and information about their small footprint TTS technology, aimed at handheld and mobile devices. Both Cepstral and Rhetorical, the technology under Appen's rVoice, use unit selection technology based on Festival, as discussed in in Steve Cassidy's TechFocus column in this issue.

## Speech on the Street

In the last issue of LT Update, we pointed to a number of voice recognition apps that are out there in use in Australia today: for example, Telstra's Directory Assistance service, VeCommerce's various TAB systems, and ESR's taxi booking service. The last six months have seen a few more systems enter the production phase. As part of the LT Update online resource base we intend to build a catalog of up-and-running speech systems in Australia and New Zealand: let us know of any you're aware of, and what you think of them, by emailing [ltupdate@ics.mq.edu.au](mailto:ltupdate@ics.mq.edu.au).

## Back to School

Macquarie University's Language Technology program is unique in Australia: see <http://www.clt.mq.edu.au/Teaching>. But there are, of course, many other places all over the world where you can undertake courses in this area. Sponsored by the field's peak international body, the Association for Computational Linguistics, the Centre for Language Technology is hosting a web-based survey of teaching in natural language processing that aims to capture useful information about courses worldwide. Check out the listings at the web site: <http://www.clt.mq.edu.au/survey>.

## A Big Hello to VeCommerce

VeCommerce is a global leader in the provision of natural language speech recognition (NLSR) and voice enabled e-commerce solutions, with over 20 years experience in the communications and call processing industries. VeCommerce's clients include the Australian Tax Office, TAB Queensland, TAB Limited, TABCORP, TOTETAS,

### 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](mailto:ltupdate@ics.mq.edu.au).*

## Upcoming Events

### National

- ANLP02, Australasian Natural Language Processing Workshop: 2 December 2002, Canberra. <http://www.clt.mq.edu.au/Events/Conferences/anlp2002/>.
- SST 2002, 9th Australian International Conference on Speech Science and Technology: 2-5 December 2002, University of Melbourne. <http://www.conferences.unimelb.edu.au/SST/>.
- AI'02, 15th Australian Joint Conference on Artificial Intelligence: 2-6 December 2002, Canberra. <http://www.cs.adfa.edu.au/~abbass/AI02/>.

### International

- ESSLLI02, 14th European Summer School in Logic, Language and Information: 5-16 August 2002, Trento, Italy. <http://www.esslli2002.it>.
- Coling2002, 19th International Conference on Computational Linguistics: 24 August - 1 September 2002, Taipei, Taiwan. <http://www.coling2002.sinica.edu.tw>.
- EDILOG 2002, Sixth Workshop on the Semantics and Pragmatics of Dialogue: 4-6 September 2002, Edinburgh University, UK. <http://www.ltg.ed.ac.uk/edilog/>.
- ICSLP2002, 7th International Conference on Spoken Language Processing: 16-20 September 2002, Denver, Colorado, USA. <http://www.icslp2002.org>.
- VoiceXML Planet and Expo: 26-27 September 2002, Omni Parker House, Boston, MA, USA. <http://www.intmediaevents.com/vxml/fall02/index.html>.
- LangTech 2002: 26-27 September 2002, Berlin, Germany. <http://www.langtech.org>.
- AMTA-2002, Association for Machine Translation in the Americas Conference: 8-12 October 2002, Tiburon, California. <http://www.amtaweb.org/AMTA2002/>.
- 4th IEEE International Conference on Multimodal Interfaces: 14-16 October 2002, Pittsburgh, PA. <http://www.is.cs.cmu.edu/icmi>.
- ISCA workshop on Advanced ASR for Telecom Applications: 27-29 November 2002, Palais des Papes, Avignon, France. <http://lands.let.kun.nl/SMADA>.
- KBCS 2002, International Conference on Knowledge Based Computer Systems: 18-21 December 2002, Mumbai, India. <http://www.ncst.ernet.in/kbcs2002>.

ACTTAB, Motor Accidents Authority, Ministry of Social Development (NZ), Auckland Co-op Taxis, Magna Entertainment Corporation (USA), Ladbrokes (UK) and Pizza Hut. We're very pleased to welcome VeCommerce as a supporting subscriber to LT Update. VeCommerce's management and operations are situated in Lane Cove, in northern Sydney. Check them out on the web at [www.vecommerce.com.au](http://www.vecommerce.com.au).

## 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 this site and you'll also find there web links to all the items mentioned in this issue as well as pointers to further resources. 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](mailto:ltupdate@ics.mq.edu.au).



# LT update

what's happening in speech  
and language technology in Australasia

## CONTENTS

- Welcome
- Tech Focus:  
Getting Computers to Talk  
Like You and Me
- G'Day Mate
- Tiny Voices
- Speech on the Street
- Back to School
- Hello VeCommerce

## Welcome to the second 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 periodic hard and soft copy publication 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 focus on text-to-speech (TTS) technologies. If what's here piques your interest, you can find out more via the links at our website: visit <http://www.clt.mq.edu.au/LTUpdate>.

### G'Day Mate

Not very long ago, any text-to-speech system (see this issue's TechFocus column) you heard in this part of the world sounded like it had a US or UK passport. Not any longer: in March, Appen launched the world's first text-to-speech technology with an Australian accent. The voice uses Rhetorical's world class core speech synthesis solution, rVoice. Appen initiated development of the product in response

## Tech Focus: Getting Computers to Talk Like You and Me

*In each issue of LT Update, we bring you a brief primer on an important area of speech and language technologies. In this issue, Steve Cassidy provides an overview of text-to-speech technologies.*

Getting computers to speak has never been an easy task. The careful inflections of HAL in 2001 stand in sharp contrast to the robotic voices we've all heard from even recent consumer grade speech synthesisers. A reaction to this has been the unwillingness of the commercial world to make use of synthetic speech in voice user interfaces; instead the majority of applications use pre-recorded speech. However, modern speech synthesis techniques can provide high quality speech output, especially if the things it needs to say are limited.

Speech synthesisers work in one of two ways: either speech signals are generated from a detailed model of human speech production, or pre-recorded speech segments are pasted together. While the model-based approach might one day produce better results, it is the cut-and-paste approach that currently gives the most intelligible output. Different sized segments can be used, but most synthesisers use phonemes (the basic building block of speech), since these allow any sentence to be generated. In order to produce natural sounding speech, the synthesiser separates out the information in the recorded signal that encodes what is being said (the phonemes) from how it is said (the tune or intonation pattern). To synthesise a sentence from text, the system first needs to work out what phonemes are needed and what intonational pattern is appropriate (is this a question? which word should be emphasised?). The pre-recorded phonemes can then be pasted together (using complex smoothing techniques so that you can't hear the joins) and combined with a new intona-

tion pattern to give the finished output.

The usual method of finding the segments to paste together is to have someone sit down and read out hundreds of nonsense words and extract diphones from these (diphones are pairs of phonemes and are used to better capture the way that one phoneme blends into the next). Another way of finding segments is to use a collection of more natural recordings and find parts of these recordings which match what's being synthesised. This method, called unit selection, can be used in limited domain synthesis where only a small set of utterances are required (for example, telling the time or reporting the weather). In limited domain synthesis, a speaker records utterances containing all the sentence patterns and all the words required and the synthesiser selects words from these for re-synthesis as required. This is almost like the cut and paste approach used in current voice interfaces, but because the phonemes and the intonation are treated separately, the result is much more natural sounding.

If you want to experiment with speech synthesis then look at the Festival research synthesiser [<http://www.cstr.ed.ac.uk/projects/festival/>] and in particular the FestVox project [<http://www.festvox.org/>] which provides tools for building new voices and includes some demos of limited domain synthesis [<http://www.festvox.org/ldom/index.html>]. There are many demonstrations of speech synthesisers available on the web: a great collection is available online as part of Sami Lemmetty's Masters thesis from Helsinki University of Technology [<http://www.acoustics.hut.fi/~slemmet/dippa/appa.html>].

Contact Steve Cassidy at [steve.cassidy@mq.edu.au](mailto:steve.cassidy@mq.edu.au).

to the dissatisfaction of Australian customers with American or British TTS products and robotic sounding voices. The Australian rVoice has been produced from recordings of an Australian female: check out Appen's website at <http://www.appen.com.au> for a demo.