

APLICACIONES DE LAS TECNOLOGÍAS DEL HABLA EN SISTEMAS CALL Y CAPT

Néstor Becerra Yoma

Laboratorio de Procesamiento y Transmisión de Voz
Departamento de Ingeniería Eléctrica
Universidad de Chile

RESUMEN

Se ha observado últimamente un gran interés en la comunidad internacional por el potencial de tecnologías de voz en las aplicaciones relacionadas con educación tales como CALL (Computer Aided Language Learning). CALL y CAPT (Computer Aided Pronunciation Training), en particular, se pueden considerar como frameworks interesantes para aplicar de modo masivo tecnología del habla. Sistemas CALL ofrecen varias ventajas sobre los métodos convencionales de clases presenciales con profesor: las lecciones se pueden preparar ad-hoc a una clase o a cada estudiante; los estudiantes, a su vez, pueden practicar ejercicios desde sus casas, laboratorio o en cualquier otro lugar en condiciones menos estresantes y embarazosas que en frente del profesor y de otros alumnos; aquellos estudiantes con dificultades de aprendizaje pueden disponer de apoyo extra para estudiar y practicar de modo más interactivo y dinámico que simplemente con un libro; y, finalmente, el problema de baja penetración en varias regiones de profesores debidamente entrenados para enseñar un segundo idioma es aliviado. Además, las tecnologías de voz tienen el potencial de proveer una retro-alimentación adecuada para corregir errores sin la necesidad imperativa de asistencia humana. Esta motivación es de especial importancia para motivar a practicar y aprender. Sin embargo, estas tecnologías sus propias limitaciones y estrategias del tipo “plug-and-play” tienden a fallar en nuestro campo.

En esta charla se discutirá el estado de avance de tecnologías CAPT, y se describirá el diseño y puesta en marcha de un sistema distribuido en Internet para la enseñanza de inglés como segundo idioma en Chile. El sistema usa la tecnología de reconocimiento de voz, basada en HMM, para la evaluación de calidad de pronunciación, y para dar respuestas por voz en actividades de comprensión de texto y de asociación de palabras a significados. La evaluación de entonación se implementa separada de la fonética utilizando estimación de pitch y alineamiento no lineal entre la elocución de test y la de referencia. La plataforma también ofrece actividades de dictado mediante las cuales el alumno debe transcribir palabras y frases. Algoritmos de programación dinámica permiten dar una nota en función del número de errores. Es interesante destacar que la plataforma propuesta puede ser vista como una etapa hacia IALL (Internet Aided Language Learning) una vez que el servicio se ofrece a través de la Internet y todo el procesamiento se realiza de modo centralizado. Finalmente, se presentan resultados preliminares de experimentos de usabilidad realizados con alumnos de un colegio público de Santiago.

Néstor Becerra-Yoma was born in Santiago, Chile, on September 15th, 1964. He received the Ph.D. degree from the University of Edinburgh, UK, and the M.Sc. and B.Sc. degrees from UNICAMP (Campinas State University), Sao Paulo, Brazil, all of them in Electrical Engineering, in 1998, 1993 and 1986, respectively. In 1998 and 1999, he was a post-doc researcher at UNICAMP and a full-time professor at Mackenzie University in Sao Paulo, Brazil. From 2000 to 2002, he was an Assistant Professor at the Department of Electrical Engineering, Universidad de Chile, in Santiago, where he is currently lecturing on telecommunications and speech processing, and working on robust speech recognition/speaker verification, dialogue systems and voice over IP. At the Universidad de Chile he has set up the Speech Processing and Transmission Laboratory (LPTV, Laboratorio de Procesamiento y Transmisión de Voz) to study speech technology applications on the Internet, education and telephone line.

Dr. Becerra-Yoma has been an Associate Professor since 2003 and is the author of 17 international journal articles and 30 conference papers. It is worth highlighting that almost all his publications are co-authored by his students. He has been the PI of two Fondef project (US\$ 450.000), three Fondecyt projects (US\$ 105.000) and two international cooperation projects. All these projects have been successfully executed, and have already finished or about to end. His results have got plenty of national press coverage in newspapers: deployment in schools of prototypes for English teaching using speech recognition and multimedia technology; and, speech recognition and vocal print technology for applications on telephony. He has supervised five M.Sc. theses and over 25

undergraduate dissertations in the last 7 years. He is also currently supervising 3 PhD students. Dr. Becerra-Yoma coordinated the commission that created the PhD program in Electrical Engineering at Universidad de Chile.

Dr. Becerra-Yoma was co-founder and elected the first chairman of the ISCA (International Speech Communication Association) Special Interest Group on Iberian Languages that already has over 100 members (professors, researchers and students) from Spain, Portugal, USA, Germany and Latin America. He has also been a chairman in several international conferences. He has been continuously invited to review papers in the most important international journal and conferences in speech and signal processing and is one of the guest editors of the special issue on Iberian languages in Speech Communications (Elsevier).

In the framework of his Fondef projects, Dr. Becerra-Yoma has organized several outreach events such as workshops, and ceremonies to officially launch the project and to deploy prototypes in schools. He has supervised the R&D activities that led to pre-competitive and competitive prototypes that are currently being evaluated by industry partners. In 2005, He got a 3rd place out of 20 proposals in a national contest for patenting.

His research interests include speech processing for telephony and education, real time Internet protocols, QoS, and usability evaluation of interfaces. Professor Becerra-Yoma is a member of the Institution of the Electrical and Electronic Engineers, and the International Speech Communication Association.