

GRAND CHALLENGES IN SPEECH RECOGNITION

Alex Acero

Microsoft Research.
Redmond, WA USA.

RESUMEN

Today, users can dial a number by simply speaking the name of the person they would like to contact, or find out whether a flight by calling a number and talking to a machine. But most people still haven't used speech technology at all or use it only infrequently, and those who run into an automated call center often request an operator or simply hang up. In the future, speech technology will be a mainstream way of interaction with software, devices and services, but advances in the state-of-the-art will be needed if we are to achieve that vision. Some of the challenges include making speech recognizers more robust to background noise and varying contexts, and designing user interfaces that are efficient and natural. In this talk, I will describe those grand challenges and some of the progress we have made at Microsoft Research. To combat background noise I will describe microphone arrays, as well as bone-conducting microphones that capture the vibration of the user's skin. I will also describe our efforts in building natural language understanding interfaces, and show tools for rapid application development. Finally I will illustrate the talk with examples of how to build speech interfaces, including multimodality for handheld devices.

Biografía: Alex Acero

Dr. Acero is Research Area Manager at Microsoft Research, overseeing natural language processing, communication, multimedia, and speech technologies. He joined Microsoft Research, Redmond, in 1994. He became Senior Researcher in 1996, manager of the speech research group in 2000, and Research Area Manager in 2005. Prior to Microsoft, Dr. Acero worked in Apple Computer's Advanced Technology Group, and Telefonica I+D. Dr. Acero is currently an affiliate Professor of Electrical Engineering at University of Washington. Dr. Acero is author of the books *Acoustical and Environmental Robustness in Automatic Speech Recognition* (Kluwer, 1993) and *Spoken Language Processing* (Prentice Hall, 2001), has written invited chapters in 3 edited books and over 120 technical papers. He holds 19 US patents. His research interests include speech recognition, synthesis and enhancement, speech denoising, language modeling, spoken language systems, statistical methods and machine learning, multimedia signal processing, and multimodal human-computer interaction. Dr. Acero is a Fellow of IEEE and 2006 Distinguished Lecturer for the IEEE Signal Processing Society. He was member of the board of governors of the IEEE Signal Processing Society between 2003 and 2005. Dr. Acero served on the Speech Technical Committee of the IEEE Signal Processing Society between 1996 and 2002, chairing the committee in 2000-2002. He was Publications Chair of ICASSP98, Sponsorship Chair of the 1999 IEEE Workshop on Automatic Speech Recognition and Understanding, and General Co-Chair of the 2001 IEEE Workshop on Automatic Speech Recognition and Understanding. He's served as Associate Editor for *Signal Processing Letters* and is presently Associate Editor for *IEEE Transactions of Speech and Audio Processing* and member of the editorial board of *Computer Speech and Language*. Alex Acero received a Master's degree from the Polytechnic University of Madrid in 1985, another Master's degree from Rice University in 1987, and a Ph.D. degree from Carnegie Mellon University in 1990, all in Electrical Engineering.