# Speech Technology: from Research to the Industry of Human-Machine Communication

**Roberto Pieraccini**
SpeechCycle
26 Broadway, 11th Floor
New York, NY 10004
roberto@speechcycle.com

## 1 Introduction

This tutorial is about the evolution of speech technology from research to a mature industry. Today, spoken language communication with computers is becoming part of everyday life. Thousands of interactive applications using spoken language technology— known also as "conversational machines"—are only phone calls away, allowing millions of users each day to access information, perform transactions, and get help. Speech recognition, language understanding, text-to-speech synthesis, machine learning, and dialog management enabled this revolution after more than 50 years of research. The industry of speech continues to mature with its evolving standards, platforms, architectures, and business models within different sectors of the market.

## 2 Content Overview

In this tutorial I will briefly trace the history of speech technology, with a special focus on speech recognition and spoken language understanding, from the early attempts to today's commercial deployments. I will summarily describe the most successful ideas and algorithms that brought to today's technology. I will discuss the struggle for ever increasing performance, the importance of data for training and evaluation, and the role played by government funded projects in creating effective evaluation benchmarks. I will then describe the birth of the speech industry in the mid 1990s, with the role played by the Voice User Interface and dialog engineering disciplines in bringing speech recognition from a laboratory "accuracy challenge" to an enabler of usable interfaces. I will describe the rising of standards (such as VoiceXML, SRGS, SSML, etc.) and their importance in the growth of the market. I will proceed with an overview of the current architectures and processes utilized for creating commercial spoken dialog systems, and will provide several case studies of the use of speech technology. I will conclude with a discussion on the current open problems and challenges.

The tutorial duration will be of about 3 hours with a short break. Several audio and video samples will be shown during the tutorial. The tutorial is directed to a general HLT audience with no prior knowledge of speech technology.

## 3 Tutorial Outline

- What is speech and why it is difficult to recognize it.
- The history of speech recognition from the early attempts to Hidden Markov Models
- The struggle for performance and the importance of data
- Spoken language understanding and dialog
- The birth of the "spoken dialog" industry
- Industrial standards and architectures
- Case studies
- Open issues and future research

## References

Pieraccini, R. Huerta, J., *Where do we go from here? Research and Commercial Spoken Dialog Systems*, Proc. of 6th SIGdial Workshop on Discourse and Dialog, Lisbon, Portugal, 2-3 September, 2005. pp. 1-10

Pieraccini R., Lubensky, D., *Spoken Language Communication with Machines: the Long and Winding Road from research to Business*, in M. Ali and F. Esposito (Eds) : IEA/AIE 2005, LNAI 3533, pp 6-15, 2005, Springer-Verlag..



*Tutorial Abstracts of ACL-08: HLT*, page 5,
Columbus, Ohio, USA, June 2008. 