December 2001

Speech-Enabled Information Systems: The Next Frontier

Alexander Hars
University of Southern California

SPEECH-ENABLED INFORMATION SYSTEMS: THE NEXT FRONTIER

Alexander Hars
Marshall School of Business
University of Southern California
hrs@bus.usc.edu

Abstract

Speech technologies are coming of age. They are applied in an increasing number of mobile, call-center, home and office settings. They challenge the established Graphical User Interface metaphor and promise to fundamentally alter the way humans conceptualize and interact with computers. This leads to new requirements for the development of information systems. It also provides new research issues and opportunities for the academic community.

This tutorial will provide an overview of natural language processing technologies, including speech recognition and generation, conversational interfaces, interactive voice response systems, voice user interfaces, translation, and standards including VoiceXML. The tutorial will provide an assessment of these technologies as well as live and recorded demonstrations. In addition, business applications of natural language technologies will be discussed. The third section of the tutorial will focus on business- and research issues of speech enabled applications. This includes issues of user-interface design, speech development life-cycles and of implications for process reengineering.

ather than focusing on the technical details of speech technologies, the purpose of this tutorial is to highlight the business and IS research issues that arise from speech-enabled information systems.

Introduction

In the last fifty years much research has been dedicated to speech technologies. While the problem has turned out to be significantly harder than anticipated, significant progress has been made. An increasing number of speech applications are being used successfully in practice. This ranges from dictation systems used in the medical and legal professions to data entry and reporting applications in hands-off environments as well as to more intelligent interactive voice response systems. However, very few IS researchers have turned their attention towards theses technologies. This is unfortunate, because speech technologies provide an abundance of research issues for academics with information systems background (and not just for computer scientists). As this tutorial will show, there is much opportunity to explore implementation strategies for speech-enabled systems, performance metrics, business models, strategies for end-user interaction and interfaces, organizational implications and business models.

In the first part of the tutorial the main components of current speech technology will be discussed. This includes the components of conversational interfaces shown in Figure 1. Demonstrations of leading-edge speech recognition and speech synthesis software will be provided. Various approaches to parsing and their limitations will be discussed. A key problem continues to be the representation of meaning and common-sense knowledge. Approaches to solve this problem will be compared. We will also discuss information extraction techniques which allow summarization and categorization of written documents. Furthermore, we will briefly review the thorniest speech technology problem: machine translation. In addition, we provide an overview over research resources for speech technologies.

In the second part of this tutorial, we will review business applications of speech technologies. We will begin with speech-enabled interactive voice response systems that are being used in many call centers. Experiences from Schwab and American Airlines will

be discussed and technologies by Nuance and General Magic will be demonstrated. In addition, VoiceXML, an emerging standard for speech-enabled IVR systems will be presented. Next, other classical speech applications in medical, legal and mobile environments will be analyzed. Subsequently, we will focus on emerging applications of speech technologies that include conversation systems being used in online store fronts (e.g. Buy.com/Soliloquy), interactive learning systems and demonstrate an interactive system for requirements elicitation (Acapulco/Questions!). Finally we will explore innovative speech applications for m-commerce.

In the last part of this tutorial business issues and research implications will be analyzed. This includes issues relating to the adoption of speech technology: Acceptance, performance metrics and feasibility analysis. Equally important is a better understanding of the speech systems life cycle. For example, speech systems require large maintenance efforts and they present considerable challenges for error and exception handling. The design of speech interfaces also provides interesting research issues: In current GUI applications, the computer primarily reacts to user commands. Context is established explicitly. In speech applications, the computer needs to become proactive – anticipating needs by the user. Instead of waiting for a query, it needs to take initiative. It needs to disambiguate user statements from its knowledge about the context the user is placed. Thus it requires some underlying domain knowledge. In addition, speech applications provide new opportunities for redesigning established business processes. Finally, it is necessary to examine the impact of speech technologies on the established computing paradigm. Speech technologies not only challenge assumptions about the graphical user interface as predominant interaction technology, they also change the location in which interaction occurs (and with it the processes), they change the way in which knowledge needs to be represented and may thus reduce the significance that relational databases currently play. The most important change, however, may be an increasing reliance on agent architectures and the inclusion of domain knowledge within information systems applications.

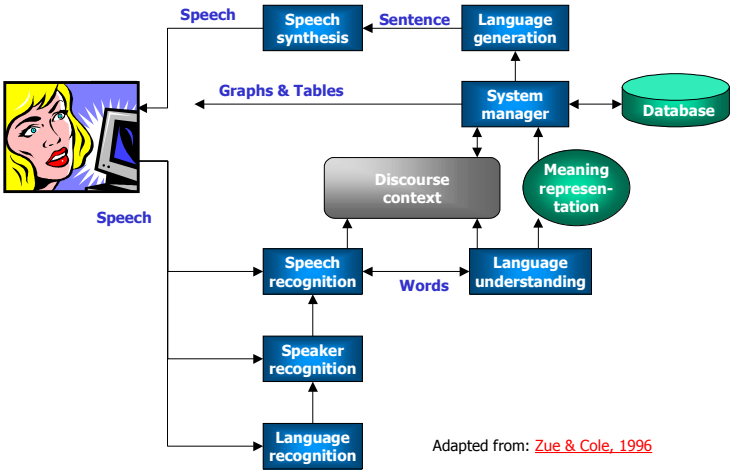

Figure 1. Technologies for Spoken Language Interfaces

Conclusion

Contrary to initial expectations, speech technologies have not yet become a high profile technology. They are far from perfect, but at the beginning of this millennium, they provide dramatic productivity improvements in many settings. We are witnessing a silent revolution that will deeply transform key aspects of information systems.

Tutorial Outline

I. State of the art in speech technologies
   - Speech/text recognition
   - Speech/text synthesis
   - Conversational interfaces
   - Machine translation
   - Information extraction
   - Encyclopedia & knowledge-bases

II. Applications of speech technologies
   - Customer service: interactive voice response
   - Expert offices (medical, legal)
   - Mobile environments and wireless
   - Help desks and IE
   - Interactive learning

III. Business issues and research implications

- Performance metrics, acceptance issues and feasibility analysis
- Speech systems life-cycle and architectures
- Business models for speech-enabled information systems
- Conversational interfaces
- Knowledge representation and ontologies
- Paradigmatic implications
- Implications for process redesign

Tool Demonstrations and Cases

- Dictation: Dragon Naturally Speaking
- IVR: Nuance, General Magic, Schwab, American Airlines
- Conversation interfaces: Alice, Brian, Maple
- Interface agents: Soliloquy: Buy.com; Dell: AskDuddley
- Knowledge bases: WordNet, OneMind
- Requirements elicitation: Acapulco, Questions!

References

Cassell, Justine (2000): Embodied conversational interface agents. In: CACM, vol.43 no.4, 70-78.

Cowie, J.; Lehnert, W. (1996): Information extraction. In: Comm. ACM, vol. 39, no.1, pp. 80-91.

Dahl, D.A.; Norton, L.M.; Scholz, K.W.: Commer-cialization of NLP technology. In: CACM, vol.43 no.11.

Dougherty, R.C. (1994): Natural language computing. Hillsdale, NJ (Erlbaum).

Hars, A.: Natural language-based data modeling: Improving validation and integration. In: J. of Database Management, vol. 9, no. 2 (Spring 1998), pp.17-25.

Hars, A.: Advancing CASE productivity by using natural language processing and computerized ontologies: The ACAPULCO system. In: Carey, J. (ed.): Proceedings of the 2nd AIS Conf., Phoenix, 1996, pp. 898-900.

Horacek, H.; Zock, M. (eds.) (1993): New concepts in natural language generation. London (Pinter).

Lewis, D.D.; Jones, K.S. (1996): Natural language processing for information retrieval. In: Comm. ACM, vol. 39, no. 1, pp. 92-101.

McCarthy, J. (1990): An example for natural language understanding and the AI problems it raises. In: Lifschitz (1990), pp. 70-76.

Nanduri, S. and Rugaber, S. (1996): Requirements validation via automated natural language parsing. Journal of MIS, Vol.12, Winter 1995-96, 9-19.

McDavid, D.W. (1996): Business language analysis for object-oriented information systems. In: IBM Systems Journal, vol. 35, no. 2, pp. 128-150.

Weinschenk, S.; Barker, D.T. (2000): Designing effective speech interfaces. New York (Wiley).

Zue, V. & Cole, R. (1996): Spoken language input: Overview. In Cole et al. (eds.): Survey of the state of the art in human language technology. http://cslu.cse.ogi.edu/HLTsurvey/ch1node3.html#SECTION11 (1/15/2001).

Zadrozny, W. et al. (2000): Enabling technologies: natural language dialogue for personalized interaction. In: CACM vol.43 no.8, 116-120.