Spoken Language Recognition and Understanding

Victor Zue and Lynette Hirschman

Spoken Language Systems Group
Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

1. PROJECT GOALS

The goal of this research is to demonstrate spoken language systems in support of interactive problem solving. The MIT spoken language system combines SUMMIT, a segment-based speech recognition system, and TINA, a probabilistic natural language system, to achieve speech understanding. The system accepts continuous speech input and handles multiple speakers without explicit speaker enrollment. It engages in interactive dialogue with the user, providing output in the form of tabular and graphical displays, as well as spoken and written responses. We have demonstrated the system on several applications, including travel planning and direction assistance; it has also been ported to several languages, including Japanese and French.

2. RECENT RESULTS

- **Improved recognition and understanding:** Reduced word error rate by over 30% through the use of improved phonetic modeling and more powerful N-gram language models; improved language understanding by 35% making use of stable corpus of annotated data; other improvements include the ability to generate a word lattice.

- **Real-time, software-only SLS system:** Developed near (1.5 times) real-time software only version of SUMMIT, using MFCC and fast match in the mixture Gaussian computation, running on a DEC Alpha or an HP735 workstation.

- **Evaluation of interactive dialogue:** Continued study of interactive dialogue, focusing on error detection and recovery issues; supported multi-site logfile evaluation through distribution of portable logfile evaluation software and instructions.

- **On-line ATIS:** Applied spoken language technology to access on-line dynamic air travel system via Compuserve; the demonstration system, extending the MIT ATIS system, provides an interactive language-based interface to find flights, make reservations and show seating assignments.

- **Multi-lingual VOYAGER:** Ported SUMMIT and TINA to Japanese, to create a speaker-independent bilingual VOYAGER; English and Japanese use the same semantic frame representation and the generation mechanism is modular and language-independent, supporting a system with independently toggled input and output languages.

- **Support to DARPA SLS community:** Chaired the ISAT Study Group on *Multi-Modal Language-Based Systems*; continued to chair MADCOW coordinating multi-site data collection, including introduction of experimental end-to-end evaluation; chaired first Spoken Language Technology Workshop at MIT, Jan. 20-22, 1993.

3. FUTURE PLANS

- **Large vocabulary spoken language systems:** Explore realistic large vocabulary spoken language applications, (e.g., on-line air travel planning), including issues of system portability and language-based interface design.

- **Multilingual knowledge-base access:** Use a uniform language-independent semantic frame to support extensions of VOYAGER and ATIS to other (more inflected) languages, e.g., French, German, Italian, and Spanish.

- **Interfacing speech and language:** Investigate loosely and tightly coupled integration, using word lattice and TINA-2's layered bigram model.

- **Dialogue modeling:** Incorporate dialogue state-specific language models to improve recognition in interactive dialogue, collect and study data on human-human interactive problem-solving, and explore alternative generation and partial understanding strategies.

- **Language modeling:** Investigate low-perplexity language models and the capture of higher level information, e.g., semantic class, phrase level information, and automatic grammar acquisition.