

## **Acoustic-Phonetics Based Speech Recognition**

Victor W. Zue

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

### *OBJECTIVE:*

The objective of this project is to develop a robust and high-performance speech recognition system using a segment-based approach to phonetic recognition. The recognition system will eventually be integrated with natural language processing to achieve spoken language understanding.

### *SUMMARY OF ACCOMPLISHMENTS:*

- Developed a phonetic recognition front-end and achieved 77% and 71% classification accuracy under speaker-dependent and -independent conditions, respectively, using a set of 38 context-independent models.
- Collaborated with researchers at SRI in the development of the MISTRI system, making explicit use of acoustic-phonetic and phonological knowledge.
- Developed the SUMMIT speech recognition system that incorporates auditory modelling and explicit segmentation, and achieved a speaker-independent accuracy of 87% on the DARPA 1000-word Resource Management task using 75 phoneme models.
- Developed probabilistic natural language system, TINA, and achieved a test-set coverage of 78% with perplexity of 42 for the Resource Management task.
- Transcribed all 6300 sentences for the TIMIT database.
- Developed a set of research tools for the DARPA speech research community in order to facilitate data collection, parameter computation, statistical analysis, and speech synthesis.

### *PLANS:*

- Improve the speech recognition performance by incorporating context-dependency in phoneme modelling.
- Integrate TINA into SUMMIT in order to develop spoken language understanding capabilities.
- Develop a back-end on the task of a Knowledgeable Navigator, and integrate it with the spoken language system.
- Begin hardware development, such that the system will soon be able to execute in near real-time.