

SRI International, Speech Research Program, Menlo Park, CA
Principal Investigators: Jared Bernstein and Hy Murveit

Objectives:

SRI's Speech Research Program focuses on developing useful speech-based systems and components for enhanced and efficient communication between human users and machines. The success of such development depends on accurate modeling of human speech and language, and the application of these models in system designs. SRI's Speech Research Program, therefore, pursues both empirical and theoretic research to gain a more comprehensive understanding of the acoustic, phonological, prosodic, lexical and syntactic nature of speech and language. SRI works on system internals such as improved performance of component technologies, as well as system-level design such as appropriate architectures and human factors solutions. Our specific technical goal is the tight integration of speech recognition and natural language understanding to create real-time systems for interactive problem solving.

Recent Accomplishments:

- SRI has developed the DECIPHER speaker-independent speech recognition system, a hidden Markov model (HMM)-based system that achieves state-of-the-art recognition performance through accurate modeling of phonetic and phonological detail.
- SRI has completed basic studies of the range and structure of variability in the pronunciation of English. In particular, SRI has shown significant differences between read speech and the spontaneous speech observed during interactive problem solving.
- SRI has designed a hardware architecture for real-time recognition of continuous speech for vocabularies as large as 20,000 words. SRI is currently implementing a subset of that architecture as a prototype accelerator with several special-purpose integrated circuits.

1989-90 Plans:

- **Hardware:** Complete the implementation of a prototype accelerator for real-time continuous speech recognition. The accelerator will use HMMs and finite state grammars to recognize a 3000-word vocabulary. Fabricate two of these accelerators for use in related sponsored research at other sites. Begin the design and fabrication of the next generation of this hardware.
- **Integration of Speech Recognition and Natural Language Understanding:** Develop a computationally efficient natural language parser that incrementally generates a state transition network that can be used in place of a finite state grammar in an HMM-based speech recognizer.
- **Spoken Language System:** Design and implement a spoken language interface for interactive problem solving in the domain of air travel planning.
- **Speech Synthesis:** Improve the potential efficiency and reliability of voice displays by the synthesis of very distinct talker identities and the synthesis of audibly different urgency levels in a meaning-to-speech generation system.
- **Neural-Net-Based Speech Recognition:** Compare the relative effectiveness of neural net, HMM-based, and hybrid approaches for specific speech recognition processes such as feature extraction and pattern classification.