

# SRI'S REAL-TIME SPOKEN LANGUAGE SYSTEM

*Patti Price and Robert C. Moore*

SRI International  
Menlo Park, CA 94025

## PROJECT GOALS

This project involves the integration of speech and natural-language processing for spoken language systems (SLS). The goal of this project, to develop a multi-modal interface to the Official Airline Guide database, is being developed along two overlapping research and development lines: one focussed on an SLS kernel for database query, and the other on the full interactive system.

## RECENT RESULTS

- Speaker-dependent and speaker-independent demos to illustrate the combined recognition/natural language system and accompanying graphical user interfaces.
- Improved robustness through the development of a template matcher for generating database queries; the template matcher has a tunable parameter to control how much constraint is ignored, so that wrong answers can be traded with no answers.
- Implementation of a bottom-up parser for CLE-formalism grammars; the new parser is about twice as fast as our original left-corner parser, and about 17 times faster than an initial bottom-up parser.
- Exploration of two schemes to integrate the recognizer and current NL schemes: N-best recognition with a statistical grammar, and recognition guided by a probabilistic finite-state representation of the templates.
- Evaluation of SRI's NL, SLS, and speech recognition technologies. SRI's February-91 weighted sentence error rate for ATIS Class A sentences was 33.8% (NL) and 44.1% (SLS); word error rate for the Resource Management speaker independent speech recognition evaluation was 17.6% with no grammar and 4.8% with the standard word-pair grammar.
- Improvements in the CLE grammar: extended coverage of numerical expressions, ATIS domain sortal restrictions, and conjoined noun phrases.
- Implementation of tied-mixture hidden Markov models, which resulted in a 20% reduction in the word-error rate compared to the discrete-density version.
- Training of a discrete-density DECIPHER using 20,000 sentences of read and spontaneous speech from Resource Management, TIMIT, and ATIS corpora. We achieved 10% word error on the June-90 ATIS test set.
- New techniques for statistical language models: a back-off estimation algorithm, Good-Turing estimates, and interpolation of word-based grammars with class-based grammars. Current test-set perplexity ranges from 15 to 30.
- Initial implementation of fast-search recognition algorithms for near real-time recognition.
- Initial implementation of speaker-adaptation using tied-mixture codebook adaptation.
- Implementation of an HMM *reject-word* model for dealing with noises and out-of-vocabulary items in digit recognition tasks; we plan to incorporate this in our ATIS SLS.

## PLANS FOR THE COMING YEAR

- Continue data collection and analysis and expand the ATIS domain.
- Implement a new architecture to combine the advantages of the robust template matcher with the new bottom-up parser. The parser will feed individual constituents to the template matcher, even if it cannot find a complete interpretation of an utterance, and will also form the basis of a phrase-level statistical language model to guide recognition (see Jackson *et al.*, this volume).
- Continue efforts in statistical language modeling: Determine why low perplexity systems do not always improve accuracy; Develop means for making efficient use of language-model training data, and for using out-of-task training data; Combine statistical language modeling with linguistic models of syntax and semantics.
- Improve acoustic modeling performance: consistency modeling, speaker-adaptation, rejection modeling (including alternative topologies and training techniques), spontaneous speech phenomena (variations in phonology and speaking rate), corrective training; Implement a continuous-distribution version of DECIPHER.