

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST)  
(Formerly National Bureau of Standards)

Principal Investigator: David S. Pallett

Our objectives in the DARPA Spoken Language Program are: (1) to provide a central role in speech database design, development and distribution; and (2) to design, coordinate implementation of, and analyze results of performance tests.

Recent Accomplishments:

(1) Developed a uniform scoring procedure and C-language software for Benchmark Tests using the DARPA Resource Management Speech Database.

(2) Coordinated and provided test material for the DARPA Speech Recognition Benchmark Tests in March and October '87, June '88, and February '89. Early tests involved BBN, CMU, MIT/Lincoln Labs and SRI. The February '89 tests are also expected to include MIT/Zue and AT&T Bell Laboratories.

(3) Prepared the Prototype CD-ROM Version of the DARPA TIMIT Acoustic-Phonetic Speech Database. This CD-ROM contains the data for 420 talkers (4200 sentence utterances and accompanying orthographic and phonetically-based transcriptions), and represents the first use of CD-ROM technology for storage and transfer of speech databases in the United States.

(4) Formed a Spoken Language Systems Performance Evaluation Working Group whose first task is to adapt benchmark test procedures for spontaneous goal-directed speech.

Plans for the remainder of FY 1989:

(1) Analyze the results of recent DARPA Benchmark Tests and provide a summary report.

(2) Prepare a report documenting the recommendations of the Spoken Language Systems Performance Evaluation Working Group and analysis of performance measures proposed, to date, in the DARPA "Performance Tasks".

(3) Assume increased responsibility for coordination and public distribution of the DARPA-community's speech databases.

(4) Prepare a CD-ROM version of the complete (630 speaker) DARPA TIMIT Acoustic-Phonetic Database, reflecting changes in format and corrections made available after release of the Prototype CD-ROM version.