Overview of the DARPA Speech and Natural Language Workshop

Lynette Hirshman
Unisys Defense Systems

The aim of the first joint Speech and Natural Language Workshop was to bring together these two research communities, to interchange technical information, to reflect on past successes and to define future directions for Spoken Language research. The overall organization of the technical program was designed 1) to establish some common reference points by assessing the current state of the art in both Speech Recognition and Natural Language Processing; 2) to cross-educate researchers in the discipline of lesser familiarity; 3) to highlight areas of common interest, namely prosodics, spoken language systems, and development of shared resources; and 4) to present current research results in both fields.

To support these aims, the program was divided into roughly four segments.
Part 1 (Day 1) focused on the establishment of common ground between the natural language and speech groups and discussion of shared resources and performance evaluation.
Part 2 focused specifically on progress in the area of Spoken Language Systems (morning of Day 2).
Part 3 (afternoon of Day 2 and morning of Day 3) focused on research results in the separate areas of speech and natural language.
Part 4 (rest of Day 3) focused on topics of joint interest: plans for selecting appropriate spoken language applications, stories of technology transfer, and finally, future plans.

In addition, two evening sessions were held. The first was for demonstrations of running systems (including video tapes). The second was an ad hoc meeting called by George Doddington (Texas Instruments) to discuss the collection of a new speech/natural language database to drive research and evaluation. This ad hoc meeting sparked a heated but profitable exchange between the groups focused on spoken language systems and those focused on more conventional speech recognition tasks. The disagreement centered on the need to evaluate the contribution of natural language *UNDERSTANDING*, as opposed to the conventional speech recognition metric of *CORRECT TRANSCRIPTION*. The spoken language researchers pushed for definition of a task to be performed via spoken language, where metrics could be developed to measure task performance independent of the notion of verbatim transcription. By the following day, there seemed to be agreement that an application such as travel planning would provide an appropriate task for purposes of research and evaluation of spoken language systems. The domain of air traffic control was also discussed at some length.

The workshop was successful in bringing these two groups of researchers together for technical discussions. The reaction to the tutorial sessions was generally very positive, with some frustration expressed at the time constraints, which precluded exploring any topic in depth. A number of researchers also expressed frustration at holding concurrent Speech and Natural Language sessions, because this prevented them from attending both sessions. This reaction was taken as a sign of serious interest in bridging the gap between these disciplines. Overall, the workshop represented a major step towards the realization of Spoken Language Systems on the part of the DARPA-funded speech and natural language research communities.