

SPEECH AND HUMAN LANGUAGE TECHNOLOGY AT THE NAVAL RESEARCH LABORATORY

Helen M. Gigley, Ph.D.

Head, Human Computer Interaction Laboratory
Naval Research Laboratory
CODE 5530
Washington, D.C. 20375-5337

ABSTRACT

Following a brief introduction to the Naval Research Laboratory (NRL), specific on-going research in speech and human language technology will be presented. Areas of research fall across two laboratories in the Information Technology Division, the Human Computer Interaction Laboratory (HCI) and the Navy Center for Applied Research in Artificial Intelligence (NCARAI). Speech research focuses on narrowband algorithm development and uses human based intelligibility to evaluate success. Demonstrations of results will be played. Navy has been a prominent participant in evaluation of narrowband speech. Current efforts on human-human communication evaluation will provide a basis for new approaches to multi-mode interactions which include speech interfaces. Finally, a video of on-going efforts at NCARAI on the EUCALYPTUS system which include graphics and spoken language interaction will be shown.

1. WHAT IS NRL?

NRL is the Department of the Navy's corporate research laboratory. Its mission is to "conduct a broadly based multidisciplinary program of scientific research and advanced technological development directed toward maritime applications of new and improved materials, techniques, equipment, systems, and ocean, atmospheric, and space sciences and related technologies." NRL reports to the Chief of Naval Research at the Office of Naval Research.

The Laboratory began in 1923 based on a vision of Thomas Edison. It presently has both a military and civilian director. There are 4 Laboratories and 17 Divisions and 14 research sites outside the Washington D.C. area. The staff is comprised of approximately 900 Ph.D.s, 500 Master's degree holders and 700 bachelor's degree holders. Each year the Lab produces a significant number of refereed journal conference publications.

The Information Technology Division houses the two laboratories whose work will be reported in the remainder of this paper, the HCI Laboratory and the NCARAI Laboratory. The work in the HCI Lab includes: speech processing within narrowband capabilities and evaluation methods for speech and other interactive systems. The effort at NCARAI in human language technology focuses

on integrating speech with graphical or other modality interfaces. A current effort is underway to transition one of the ARPA speech technology systems to the on-going work in EUCALYPTUS.

1.1 Speech Technology in HCI

The speech technology group in the HCI Lab focuses on algorithmic development which is reaching toward communication of humanly intelligible speech at 600 bps. Current algorithms at 1200 bps are quite successful. The method relies on linear predictive coding (LPC) of the speech signal to LPC coefficients which can be independently manipulated to enhance aspects of the signal under differing conditions. Changes in pitch with all else constant and in rate of speech by slowing or speeding the presentation of the entire set of coefficients are found to be equally intelligible and without distortions that interfere with intelligibility. A sample of a conversation at different communication bandwidths further demonstrates the role of the technology. This group has recently developed, using an off-the-shelf Macintosh, a non-real-time terminal which can transmit both speech and images (including annotation) within 2400 bps. A second terminal is being developed to transmit voice and still images simultaneously but real-time.

1.2 Evaluation Methods and Technology for Speech and Interaction

The Navy has taken an active part in evaluation of speech systems by participation in the Digital Voice Consortium since its inception in 1972. The Consortium coordinates speech efforts across DoD and insures interoperability of systems as well as attempts to eliminate duplication of efforts. The evaluation group in the HCI Lab has contributed significantly to the test set used in evaluating the intelligibility of narrowband speech.

A new project which spans both the HCI Lab and the NACRE Lab in human machine dialogue has recently shifted the focus to development of new evaluation methods to study human-human interactions as a basis for developing new metaphors for human machine interaction. Using AEGIS team training data, issues of speech turn taking during decision making and length of speech and counts of phrases have been studied. This group provides expert consulting on evaluation for all levels of interactive techniques which include speech and human language uses.

1.3 Human Language Technology at NCARAI

The Interactive Systems Group at NCARAI has recently applied a SSI speech recognizer as input to the natural language interface which was added to a functioning interactive decision support tool called KOALAS. The natural input, EUCALYPTUS, assumes spoken language inputs (not just key phrases) are available. The

interactions with the system use existing graphical semantics of the domain and integrate the language capability to refer to objects which are graphically presented. There is no extension to the graphical interface. Only domain relevant semantics are available to the language processor. The system achieves an integrated graphics / natural language interaction which can be used separately or in combination. Human Language capabilities such as ellipsis, deixis, and anaphora resolution have been implemented which allow cross-modal and contextually determined reference. A video of the working system documents the functionality.

1.4 On-going Experiment in ARPA Speech Technology Transition

Within the HCI Lab work is currently underway to port the Sphinx II system to a SPARC 10. It will then replace the current EUCALYPTUS front end processor. This is a simple goal as we already have a working speech capability for the domain and only intend to substitute one "black-box" for another. We are discovering hidden constraints in the porting across hardware platforms which need to be handled explicitly if these systems are to become readily available to the computing public. This is an experiment in progress with only one of the systems. Several others will be used in similar fashion as a means of developing transition methods to integrate the results into real military applications. Whatever is discovered will contribute to the successful transition to dual-use areas as well.

2. POINTS OF CONTACT

The Section Heads for each of the research areas presented are as follows:

Speech Technology

George Kang, HCI Lab, CODE 5531

email: kang@itd.nrl.navy.mil

Evaluation Methods and Technology

Dr. Astrid Schmidt-Nielsen, HCI Lab, CODE 5532

email: schmidt@itd.nrl.navy.mil

Interactive Systems

Elaine Marsh, NCARAI, CODE 5512

email: marsh@aic.nrl.navy.mil

ACKNOWLEDGMENT

The research described is supported by the Office of Naval Research within 6.1 and 6.2 Block grants to the Naval Research Laboratory HCI and NCARAI Laboratories and by SPAWAR.