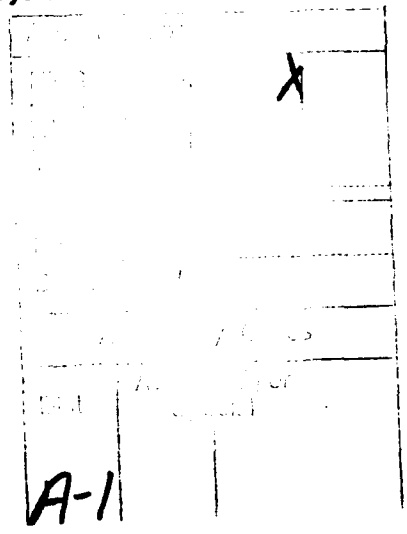
(2)

BBN's SLS ANNUAL REPORT for ONR

Principal Investigator(s): Dr. John Makhoul & Dr. Madeleine Bates
PI Institution: BBN Systems & Technologies
PI Phone Number: Dr. Makhoul 617-873-3332; Dr. Bates 617-873-3634
PI E-mail Address: makhoul@bbn.com; bates@bbn.com
Grant or Contract Title: Usable , Real-Time, Interactive Spoken Language Systems
Grant or Contract No: N00014-92-C-0035
Reporting Period: 1 October 1991 - 30 September 1992 (*)



1. Productivity measures.

Refereed papers submitted but not yet published: 0

Refereed papers published: 16

Unrefereed reports and articles: 0

Books or parts thereof submitted but not yet published: 0

Books or parts thereof published: 0

Patents filed but not yet granted: 0

Patents granted (include software copyrights): 0 patents

Invited presentations: 1

Contributed presentations: 0

Statement A per telecon
LCDR Robert Powell ONR/Code 1133
Arlington, VA 22217-5000

10/26/92 NWW

Honors received (fellowships, technical society appointments, conference committee roles, editorships, etc.): also include descriptions of the specific honors. 3 (see list in Section 4)

Prizes or awards received (Nobel, Japan, Turing, etc.): 0

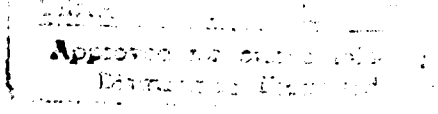
Promotions obtained: 0

Graduate students supported >= 25% of full time: 0

Post-docs supported >= 25% of full time: 0

Minorities supported (include Blacks, Hispanics, American Indians and other native Americans such as Aleuts, Pacific Islanders, etc.; do not include Asians or Asian-Americans): 0

(*) Work in the period from 1 October 1991 through 31 January 1992 was performed under Contract N00014-89-C-0008, Spoken Language Systems. Work since 1 February 1992 was performed under the contract indicated above.



Principal Investigator(s): Dr. John Makhoul & Dr. Madeleine Bates
PI Institution: BBN Systems & Technologies
PI Phone Number: Dr. Makhoul 617-873-3332; Dr. Bates 617-873-3634
PI E-mail Address: makhoul@bbn.com; bates@bbn.com
Grant or Contract Title: Usable , Real-Time, Interactive Spoken Language Systems
Grant or Contract No: N00014-92-C-0035
Reporting Period: 1 October 1991 - 30 September 1992

2. Detailed summary of technical progress.

2.1 Background and Goals

To adapt to the requirements posed by rapidly changing world-wide threats, modern military systems must harness advanced technology in ways that significantly improve operational effectiveness. Many critical systems are interactive, and the ability of users to interact with such systems could be improved by the addition of spoken language interfaces to facilitate human/machine interaction, increase productivity, and reduce training time.

To make the next significant advance in human-machine interaction for military systems, we must bridge the gap between feasible and usable. Not only must we advance the state of the art in high-accuracy speech understanding, but we must demonstrate the utility of real-time SLS in applications that are clearly relevant to the military. At BBN Systems & Technologies, we are making rapid progress toward those ends.

Our goals are to develop usable SLSs that exhibit the following advances:

1. at least an order of magnitude increase in speed, with higher accuracy;
2. a four-fold reduction in the overall understanding error rate;
3. a vocabulary of up to 10,000 words;
4. a highly interactive user interface capable of mixed initiative; dialogue, system feedback, and user corrections and additions;
5. a flexible system capable of transparently adapting to a new user;
6. a modular system, easily portable to new applications and sites;
7. a system implementable in real-time on COTS hardware.

Our approach to these goals includes:

- Accurate and fast search algorithms incorporating techniques such as: fast match, variable frame rate, cross-word modeling, and advanced speech grammars.
- Robust NL processing based on combining fragments of partially understood language, and a new hybrid method of knowledge representation that combines features of logical and frame representations.

- A novel, potentially high-payoff approach to language processing that incorporates a statistical model of meaning, which should lead to automatic knowledge acquisition in NL, coupled with high performance.
- New methods for mixed-initiative dialogue which are tolerant of common user faults, and allow for corrections and additions by the user.

2.2 Major Accomplishments

In February, 1992, BBN participated in the DARPA Speech & Natural Language Workshop. In the official results scored by NIST, BBN had the best results of all sites participating in two tests:

(1) speech recognition (9.4% word error rate over all utterances tested), and
(2) overall spoken language understanding (43.7% weighted error over all utterances tested).

At this workshop, BBN gave a demonstration of the first 1000-word, real-time, continuous, speaker independent speech recognition system implemented on an off-the-shelf workstation, without any accelerator boards.

BBN also gave the first real-time demonstration of a complete spoken language understanding system in the ATIS (Air Travel Information System) domain.

Our system continues to have very low answer inflation (overgeneration), showing the precision of our natural language understanding system.

We continued to improve the capabilities of the natural language system without adversely impacting the speed of the system.

We added a spoken output capability to our demonstration system, which provides a conversational "feel" to the system.

We continued to chair and to work closely with several DARPA-wide SLS program committees, in particular the committee that is determining the requirements for the work in continuous speech recognition (BBN chairs the CCCC) and the committee that determines the methodology for the ATIS systems (the MADCOW committee).

2.3 Research in Natural Language Processing and System Integration

BBN's research on natural language (NL) processing has concentrated on improved methods of integrating syntactic and semantic processing in an efficient manner that will also be portable to different underlying application systems. The BBN NL system, called DELPHI, has automatic facilities for handling assumptions about objects that are not fully known (which is useful for attaching probabilities and other annotations to semantic interpretations).

We have implemented a fallback mechanism to use when a syntactic and semantic analysis of the complete utterance cannot be produced; this permits the BBN system to handle ungrammatical (or mis-recognized) speech very robustly.

Our research into the appropriate integration of lexical, syntactic, and semantic knowledge has led us toward a system that is capable of using efficiently whatever types of knowledge will produce a valid interpretation of an utterance in context. Syntactic knowledge is used if it is reliable, but an utterance that is outside the scope of DELPLHI's grammar can still be understood if phrases can be recognized and combined semantically. We began work on a learnable model of semantics, which will enormously facilitate the acquisition of domain-specific information that was formerly very labor intensive to produce.

We have also experimented with several conditions to optimize the integration of the BYBLOS speech recognition system with DELPHI. The basic interface between them is the N-Best list. We determined empirically that the optimum N for our system is 5.

2.4 Research in Speech Processing

Our speech recognition work under the SLS contract falls into two main areas: fast match search algorithms and three demonstration systems.

Fast Match Search. We implemented the fast match algorithm proposed by Dragon systems. The purpose of the fast match algorithm is to return a reduced list of likely word candidates for further detailed scoring. However, we found that, while the scores produced by the algorithm were correlated with whether the word was correct, there were several instances where the correct word had a fairly low score, such that the correct word would have been eliminated from consideration, thus increasing the overall error rate. Therefore, we do not intend to use this fast match method.

Real-Time ATIS System. In cooperation with other projects at BBN, we developed a 1000-word continuous speech recognition system that runs in real-time on an off-the-shelf workstation, with no additional accelerators used. The system was demonstrated in several applications, including air traffic control and ATIS. The ATIS demonstration performed the understanding in several rapid phases. A first-pass speech recognition is performed in real-time as the user speaks. As soon as the user stops speaking the result of this 1-best recognition is printed. Then, the system takes another second to find the N-best sentence hypotheses. These are reordered using a more powerful trigram statistical language model. We found that this rescoring step reduced the word error rate on relevant sentences by about 30%. Finally, the reordered N-best list of hypotheses is sent to the BBN DELPHI language understanding component, which chooses the sentence that makes the most sense, performs the appropriate database retrieval, and displays the answer. To our knowledge, this is the first and only system that performs large-vocabulary speech recognition and understanding in real time without the aid of additional accelerators.

Demonstration System at DARPA. We delivered a workstation with several different demonstrations of speech recognition, speech understanding, and text data extraction to the contract monitor at DARPA. The speech recognition demonstrations included were on the Resource Management domain, and an Air Traffic Control application that we are developing for the FAA. The demonstration of real-time spoken language understanding in the ATIS domain using the BBN HARC system demonstrates the efficient combination of the BYBLOS speech recognition system and the Delphi natural language system. We also included a demonstration of the BBN PLUM system performing a text data extraction task. Finally, we included a facility to play examples of the speech from each of several different DARPA speech corpora.

Voice Command. We implemented a Voice Command system using the real-time recognition system. This system provides a simple mechanism for a user to associate various word sequences with specific actions. The demonstration, which runs interactively, allows the user to modify the set of phrases and actions. That is, the user can easily and rapidly add words to the vocabulary and specify their meanings in terms of actions to be performed by the system. The system was demonstrated at the Spoken Language Steering Committee meeting in August.

Principal Investigator(s): Dr. John Makhoul & Dr. Madeleine Bates
PI Institution: BBN Systems & Technologies
PI Phone Number: Dr. Makhoul 617-873-3332; Dr. Bates 617-873-3634
PI E-mail Address: makhoul@bbn.com; bates@bbn.com
Grant or Contract Title: Usable , Real-Time, Interactive Spoken Language Systems
Grant or Contract No: N00014-92-C-0035
Reporting Period: 1 October 1991 - 30 September 1992

3. Lists of publications, presentations and reports.

Steve Austin, John Makhoul,and Richard Schwartz, BBN; G. Zavaliagkos, Northeastern University, "Speech Recognition Using Segmental Neural Nets", presented at ICASSP, 1992.

Steve Austin, G. Zavaliagkos, John Makhoul and Richard Schwartz, "Improving State-of-the-Art Continuous Speech Recognition Systems Using the N-Best Paradigm with Neural Networks", presented at the 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Madeleine Bates, et al, "Commercial, Isolted Word Speech REcognition Integrated with NL Processessing for Intelligence Applications", *IEEE 1992 C3I Technology and Applications Conference*, Rome, NY, June 1-4, 1992.

Madeleine Bates and S. Boisen, "A Developing Methodology for the Evaluation of Spoken Language Systems", Workshop on Evaluation of Natural Language Processing Systems, *29th Annual Meeting of the Association for Computational Linguistics*, Berkeley, CA, June 19-21, 1991.

Robert Bobrow, Robert Ingria, and David Stallard, "Syntactic/Semantic Coupling in the BBN DELPHI System", 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Herbert Gish, "A Minimum Classification Error, Maximum Likelihood, Neural Network, presented at ICASSP, 1992.

Robert Ingria, "DARPA Common Lexicon Progress Report", presented at the 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Owen Kimball, Mari Ostendord and Robin Rohlicek, "Recognition Using Classification and Segmentation Scoring", presented at the 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Francis Kubala, et al, "BBN BYBLOS and HARC February 1992 ATIS Benchmark Results", 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Kenney Ng, Herbert Gish and Robin Rohlicek, "Robust Mapping of noisy Speech Parameters for HMM Work Spotting, presented at ICASSP, 1992.

MADCOW (Madeleine Bates was a member of this committee), "Multi-Site Data Collection for A Spoken Language Corpus", presented at the 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Robin Rohlicek, BBN; Mari Ostendorf, Boston University, "A Bayesian Approach to Sepaker Adaptation for the Stochastic Segment Model, presented at ICASSP, 1992.

Robin Rohlicek, Damaris Ayuso, Madeleine Bates, Rusty Bobrow, Albert Boulanger, Herbert Gish, Philippe Jeanrenaud, Marie Meteer and M. Siu, "Gisting Conversational Speech", presented at ICASSP, 1992.

Richard Schwartz, Steve Austin, Francis Kubala, John Makhoul, Long Nguyen, Paul Placeway; George Zavaliagkos, Northeastern University, "New Uses for the N-Best Sentence Hypotheses Within the Byblos Speech Recognition System", presented at ICASSP, 1992.

Man-Hung Siu, George Yu and Herbert Gish, "An Unsupervised, Sequential, Learning Algorithm for the Segmentation of Speech Waveforms with Multiple Speakers", presented at ICASSP, 1992.

David Stallard and Robert Bobrow, "Fragment Processing in the DELPHI System", 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

BBN also presented demonstrations of our real-time ATIS system and real time Speech Recognition System for Air Traffic Controllers during the demonstration session at the 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992; and the BBN real time ATIS system was available in the demo room throughout the workshop.

Principal Investigator(s): Dr. John Makhoul & Dr. Madeleine Bates
PI Institution: BBN Systems & Technologies
PI Phone Number: Dr. Makhoul 617-873-3332; Dr. Bates 617-873-3634
PI E-mail Address: makhoul@bbn.com; bates@bbn.com
Grant or Contract Title: Usable , Real-Time, Interactive Spoken Language Systems
Grant or Contract No: N00014-92-C-0035
Reporting Period: 1 October 1991 - 30 September 1992

4. Transitions and DoD interactions.

Madeleine Bates co-chaired the Applied Natural Language Processing Conference in Trento, Italy in April, 1992.

Madeleine Bates is chairing the Program Committee for the 6th DARPA S&NL Workshop.

Madeleine Bates is a member of the Program Committee for the 9th IEEE Conference on Artificial Intelligence for Application, to be held in Orlando, Florida in March, 1993

Robert Ingria chaired a meeting of the DARPA Common Lexicon Working Group at the 5th DARPA Speech & NL Workshop at Arden House, February 23-26, 1992.

Francis Kubala is chairing the DARPA CCCC committee.

Madeleine Bates, Robert Bobrow, and Robert Ingria attended the 30th Annual Meeting of the Association for Computational Linguistics at the University of Delaware, Newark, Delaware, June 28-July 2, 1992.

Madeleine Bates was co-chair of a session on the Integration of Natural Language and Speech, for the IEEE 1992 C3I Technology and Applications Conference, SUNY Institute of Technology, Utica/Rome New York, June 1-4, 1992.

Robert Ingria attended a meeting of the Working Group on Syntactic Evaluation at the 30th Annual Meeting of the Association for Computational Linguistics at the University of Delaware, Newark, Delaware, June 28-July 2, 1992.

Principal Investigator(s): Dr. John Makhoul & Dr. Madeleine Bates
PI Institution: BBN Systems & Technologies
PI Phone Number: Dr. Makhoul 617-873-3332; Dr. Bates 617-873-3634
PI E-mail Address: makhoul@bbn.com; bates@bbn.com
Grant or Contract Title: Usable , Real-Time, Interactive Spoken Language Systems
Grant or Contract No: N00014-92-C-0035
Reporting Period: 1 October 1991 - 30 September 1992

5. Software and hardware prototypes.

Because of the success of our demonstrations at the DARPA workshop, our program manager at DARPA, Mr. Charles Wayne, requested that we duplicate our real-time system at DARPA. We delivered an SGI Indigo computer to DARPA, with a menu-driven, real-time demonstration system capable of showing several systems (including the BBN ATIS System, the BBN Air Traffic Controller Speech Recognition System, real time speech recognition in the Resource Management domain, and a system that integrates speech recognition in various command and control demonstrations).

We also provided a videotape to DARPA showing various capabilities of some of our real-time systems.