

December 2001

Military Applications of Natural Language Processing and Software

James Rodger

Indiana University of Pennsylvania

Parag Pendharkar

Penn State Capital College

David Paper

Utah State University

Tamara Trank

MSC, USNR

Follow this and additional works at: <http://aisel.aisnet.org/amcis2001>

Recommended Citation

Rodger, James; Pendharkar, Parag; Paper, David; and Trank, Tamara, "Military Applications of Natural Language Processing and Software" (2001). *AMCIS 2001 Proceedings*. 233.

<http://aisel.aisnet.org/amcis2001/233>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISEL). It has been accepted for inclusion in AMCIS 2001 Proceedings by an authorized administrator of AIS Electronic Library (AISEL). For more information, please contact elibrary@aisnet.org.

MILITARY APPLICATIONS OF NATURAL LANGUAGE PROCESSING AND SOFTWARE

James A. Rodger
Indiana University of Pennsylvania
jrodger@grove.iup.edu

Parag C. Pendharkar
Penn State Capital College
pxp19@psu.edu

David J. Paper
Utah State University
dpaper@b202.usu.edu

Lt. Tamara V. Trank, MSC
USNR, NHRC
trank@nhrc.navy.mil

Abstract

The goal of this research is to enhance Force Health Protection and improve medical readiness by applying voice interactive technology to environmental and clinical surveillance activities aboard U.S. Navy ships. In support of this goal, the project team has developed a lightweight, wearable computing device with voice interactive capability, enabling environmental surveillance to be accomplished quickly and efficiently without compromising the quality of surveillance information. To the extent possible, existing technologies were utilized in creating the device, which is capable of storing, processing, and forwarding data to a server, as well as interfacing with other systems, including Shipboard Nontactical Automated Data Processing (SNAP) Automated Medical Systems (SAMS). The voice interactive computing device includes automated user prompts, enhanced data analysis, presentation, and dissemination tools, and computer-based training in support of preventive and clinical medicine. In addition to reducing the time needed to complete inspections, the device has been designed to support local reporting requirements, improve hazardous material exposure response, and enhance command level intelligence.

Background

When computer recognition of the human voice is coupled with a natural language processing system, speech recognition by computers is made possible. By allowing data and commands to be entered into a computer without the need for typing, computer understanding of naturally spoken languages frees human hands for other tasks. Speech recognition by computers can also increase the rate of data entry, improve spelling accuracy, permit remote access to databases utilizing wireless technology, and ease access to computer systems by those who lack typing skills.

Speech Recognition Applications

Although speech recognition systems have existed for two decades, widespread usage of this technology is a recent phenomenon. As improvements have been made in accuracy, speed, portability, and operation in high noise environments, the development of speech recognition applications by the armed services, federal agencies, and private sector has increased.

One example is Boeing's testing of the Voxware system's speech recognition technology in the noisy cockpit of its Joint Strike Fighter. The Voxware system includes a rugged, lightweight, continuous-speech device that permits the operation of selected cockpit controls by voice commands alone. The objective is to free pilots from certain manual tasks and sharpen their focus on the flight environment. Another effort involving pilots has been conducted by the Naval Aerospace Medical Research Lab. Pilots must have good head/eye coordination when they shift their gaze between cockpit instruments and the outside environment. The Naval Aerospace Medical Research Lab has tested the feasibility of using speech recognition to support the measurement of these shifts and the type of coordination required to make them. In another aviation-related effort, the Federal Aviation Administration

conducted a demonstration of how voice technology supports a facilities maintenance task. A voice-activated system proved to be less time-consuming to use than the traditional paper manual approach, and study participants reported that the system was understandable, easy to control, and responsive to voice commands. Participants felt that the speech recognition system made the maintenance task easier to perform, was more efficient and effective than a paper manual, and would be better for handling large amounts of information.

Speech recognition technology is expected to play an important role in supporting real time interactive voice communication over distributed, computer data networks. The Interactive Voice Exchange Application developed by the Naval Research Lab has been able to maintain a low data rate throughput requirement and permits the use of voice communication over existing computer networks without causing a significant impact on other data communications, such as e-mail and file transfer.

Existing Problems

Though applications of speech recognition technology have been developed with increased frequency, several problems with speech recognition systems have yet to be resolved. Problems in the compatibility and integration of recognition tools with underlying computer operating systems have an impact on applications development.

Use of speech recognition technology in high noise environments also remains a challenge. Normally, for speech recognition systems to function properly, clean speech signals are required, with high signal to noise ratio and wide frequency response. A microphone system is critical in providing the speech signal required for recognition accuracy.

In high noise environments, providing a clean speech signal can be difficult. Interference, changes in the user's voice, and additive noise – such as car engine noise, background chatter, and white noise – can reduce the accuracy of speech recognition systems. In military environments, additive noise and voice changes are common. For example, in military aviation, the stress resulting from low-level flying and other difficult missions can cause a speaker's voice to change, reducing recognition accuracy. Until recently, few practical continuous speech recognizers were available. Most were difficult to build, resided on large mainframe computers, were speaker dependent, and did not operate in real time. The Voice Interactive Display (VID) developed for the U.S. Army has made progress in eliminating these disadvantages. VID was intended to reduce the bulk, weight, and set-up times of military diagnostic systems while increasing their capacity and capabilities for hands-free trouble shooting. The capabilities of VID were developed to allow communication with the supply and logistics structures within the Army's common operating environment.

This effort has matured and now demonstrates the use of VID as a tool for providing a paperless method of documentation for diagnostic and prognostic results. The VID system will culminate in the automation of maintenance supply actions. Thanks to funding by the Tank-Automotive & Armaments Command (TACOM) through the National Automotive Center (NAC), voice recognition technology and existing diagnostic tools have been integrated into a wireless configuration. The result is a hands-free interface between the operator and the Soldier's On-System Repair Tool (SPORT).

The VID system consists of a microphone, a hand-held display unit, and SPORT. With this configuration, a technician can obtain vehicle diagnostic information while navigating through an Interactive Electronic Technical Manual via voice commands. By integrating paperless documentation, human expertise, and connectivity to provide user support for vehicle maintenance, VID maximizes U.S. Army efficiency and effectiveness. VID enhances the readiness of U.S. Army forces during peacetime and provides for rapid communication to military logistics worldwide.

Setting the Stage

Encouraged by the success of the Army's VID project, the U.S. Navy launched a VID project of its own. The goal of the Navy Voice Interactive Device (NVID) project is to create a lightweight, portable computing device that uses speech recognition to enter shipboard environmental survey data into a computer database.

To ensure the health and safety of shipboard personnel, Naval health professionals (NHPs) – including environmental health officers, industrial hygienists (IHs), independent duty corpsmen (IDCs), and preventive medicine technicians (PMTs) – perform clinical activities and preventive medicine surveillance on a daily basis. In addition to providing patient care, IDCs perform water testing, heat stress surveys, pest control surveys, and routine sanitation and habitability inspections.

OPNAVINST 5100.19D, the *Navy Occupational Safety and Health Program Manual for Forces Afloat*, provides the specific guidelines for maintaining a safe and healthy work environment aboard U.S. Navy ships. Inspections performed by NHPs ensure that the guidelines are followed. The various shipboard surveys are conducted by shipboard personnel on a continual basis and by NHPs from local support facilities (Naval Medical Centers or Environmental Preventive Medical Units) when a ship is in port. Typically, NHPs enter data and findings by hand onto paper forms and later transcribe these notes into a word processor to create a finished report. The process of manual note-taking and entering data via keyboard into a computer database is time-consuming, prone to error, and difficult to accomplish in a cramped shipboard environment. To remedy these problems, the Naval Shipboard Survey Planning System (NSSPS) was developed, allowing data to be collected and transferred into portable laptop computers while a survey is being conducted.

The use of laptop computers proved to be difficult because of space constraints and the inability of inspectors to have both hands free to type during an inspection. The need was identified for a hands-free, space-saving mode of data entry that would enable NHPs to enter data or obtain instructions and information. Out of this need, the NVID project was created. NVID's goal is to replace existing, inefficient, repetitive survey procedures with a fully automated, voice interactive system for voice-activated data input. In pursuit of this goal, the NVID team has developed a lightweight, wearable, voice-interactive computer capable of capturing, storing, processing, and forwarding data to a server for easy retrieval by users.

The voice interactive data input and output capability of NVID reduces obstacles to accurate and efficient data access, reduces the time required to complete inspections and greatly improves response to hazardous material exposure incidents. NVID's voice interactive technology allows a trainee to interact with a computerized system and still have hands and eyes free to manipulate materials and negotiate his or her environment. Once entered, survey and medical encounter data can be used for local reporting requirements and command level intelligence. Improved data acquisition and transmission capabilities will allow connectivity with other systems.

Existing printed and computerized surveys are voice activated and reside on the miniaturized computing device. NVID has been designed to allow voice prompting by the survey program, as well as voice-activated, free text dictation. An enhanced microphone system permits improved signal detection in noisy shipboard environments. The NVID technology also provides voice interactive capability documenting medical encounters using the SAMS shipboard medical database. This technology should prove particularly useful in aiding active and sometimes overburdened medical providers with neither time nor, in some cases, the typing proficiency, for rapidly and accurately entering vital patient charting data.

NVID's capability to access data from SAMS enhances the NHPs' ability to identify trends and health hazard exposures. Researchers at Naval Health Research Center (NHRC) are developing a clinical data analysis tool, the Epidemiological Wizard (EPIWIZ), that extracts medical data from SAMS and generates summary reports used for detecting environmental changes and early identification of disease and injury trends. In turn, this data will be analyzed in order to identify changes that may be indicative of an exposure to a health hazard.

By integrating such data analysis tools and other emergent medical information elements with NVID's voice recognition technology, the NVID team plans to expand the ability of operational force commanders to detect disease and injury trends early, allowing quicker intervention to prevent force degradation.

NVID's capability to store, process, and forward data to a server can be a useful tool in assisting with environmental surveillance, providing rapid response to hazardous material exposure incidents, and enhancing shipboard preventive medicine. Other applications of NVID are being explored outside the armed services. One object of this effort is to investigate the feasibility of using voice interactive technology to support local occupational health, preventive medicine, and clinical care.

Literature Review

It is desirable to investigate the feasibility of using voice interactive technology to support local occupational health and preventive medicine aboard ships. Integration of existing technologies should be the intent of the design rather than new technology development. A lightweight, voice-activated computer capable of storing, processing, and forwarding data to a server shall be a useful tool to aid industrial hygienists in practicing occupational health and preventive medicine aboard ship. Users of the prototype could determine if the application is feasible and evaluate modifications and integration of NSSPS and EPIWIZ software into a Navy Voice Interactive Device (NVID) prototype. Such a system could, with further modifications, be used in a variety of other medical and clinical applications.

A basis for evaluating the NVID was provided by the National Institute of Standards and Technology (NIST). NIST has supported other federal government agencies' speech research programs, and the university and industrial research communities since 1987 (Pallet, Garfolo and Fiscus, 2000). In a 1998 broadcast news test, NIST provided participants with a test set consisting of two 1.5 hour subsets obtained from the Linguistic Data Consortium. The task associated with this material was to implement automatic speech recognition technology by determining the lowest word error rate (Herb and Schmidt, 1994; Pallett, 1999; Fiscus, 1997; Greenberg, Chang, and Hollenback, 2000). Excellent performance was achieved at several sites, both domestic and abroad (Przbocki, 1999). For example, IBM-developed systems achieved the lowest overall word error rate of 13.5%. The application of statistical significance tests indicated that the differences in performance between systems designed by IBM, the French National Laboratories' Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingenieur (LIMSI) and Cambridge University's HTK (CU-HTK) were not significant (Pallet, Garfolo, and Fiscus, 2000).

Similar findings were reported by other research (Lai, 2000). These research findings show that no significant differences exist in the comprehension of synthetic speech among five different speech-to-text engines used. Just as different speech-to-text engines produce similar results, speech comprehension can work as effectively with different languages as with English. Studies have documented both English and Mandarin word error rates of 19.3% (Fiscus, Fisher, Martin, Przybocki and Pallet, 2000). Finally, speaker segmentation has been used to locate all the boundaries between speakers in the audio signal and enables speaker normalization and adaptation techniques to be used effectively to integrate speech recognition (Bikel, Miller, Schwartz, and Weischedel, 1997).

Speech recognition technologies are in their infancy, and many problems remain to be investigated. For example, the success of speech recognition by desktop computers depends to a large extent on the integration of speech technologies with the underlying processor and operating system and the complexity and availability of tools required to deploy a system (Markowitz, 2000; Woo, 2000). The accuracy of modeling has also limited the effectiveness of speech recognition. Modeling accuracy can be improved, however, by combining feature streams with neural nets and Gaussian mixtures (Ellis, 2000). The application of knowledge-based speech analysis has also shown promise (Komissarchik, and Komissarchik, 2000).

It is assumed that for speech recognition systems to function properly, they require clean speech signals with significant signal to noise ratio and wide frequency response (Albers, 2000; Erten, Paoletti, and Salam, 2000; Sones, 2000; Wickstrom, 2000). The microphone system is critical in providing the required speech signal, and, therefore, has a direct effect on the accuracy of the speech recognition system (Andrea, 2000; Wenger, 2000). It is well known that the control of the speech recognition interface poses its own unique problems (Taylor, 2000; Gunn, 2000). The inability of people to remember verbal commands is even more of a hindrance than their inability to remember keyboard commands (Newman, 2000). The limited quality of machine speech output also affects the speech recognition interface. As man-machine interaction becomes increasingly commonplace, applications that require unlimited vocabulary speech output are demanding text-to-speech (TTS) systems that produce more human-sounding speech (Hertz, Younes, and Hoskins 2000).

Garofolo, Pallett and Fiscus (1999) point out that potential problems associated with the search and retrieval of relevant information from databases have been addressed by the Spoken Document Retrieval (SDC) community. Furthermore, standards for the probability of false alarms and miss probabilities are set forth and investigated by the Topic Detection and Tracking (TDT) program (Dottington, 1999). Decision Error Tradeoff (DET) plots are used to demonstrate the tradeoff between the miss probabilities and false alarm probabilities for a topic (Kubala, 1999). Security issues and speech verification are major voids in speech recognition technology (Gagnon, 2000). Technology for the archiving of speech is also undeveloped. It is well recognized that speech is not presently valued as an archival information source because it is impossible to locate information in large audio archives (Kubala, Colbath, Liu, Srivastava, and Makhoul, 2000).

Applications of Speech Recognition

Many applications of speech technology exist. For example, BT Adastral Park receives approximately 700 million calls each year at its UK directory assistance bureau, a significant proportion of which are for residential numbers (Gorham and Graham, 2000). Studies in such call centers have utilized live customer trials to demonstrate the technical realization of full speech automation of directory inquiries (McCarty, 2000; Miller, 2000). Performance of the service, a study of customer behavior, and an analysis of service following call-back interviews shed a favorable light on the applications of speech automation to this industry (Gorham and Graham, 2000). In another application of speech recognition technology, Haynes (2000) deployed a conversational Interactive Voice Response (IVR) system to demonstrate site-specific examples of how companies are leveraging their infrastructure investments, improving customer satisfaction and receiving quick return on investments.

Such applications demonstrate the use of speech recognition by business. The investigation of current customer needs and individual design options for accessing information utilizing speech recognition is key to gaining unique business advantages (Schalk, 2000; Prizer, Thomas and Suhm, 2000).

Internationalizing speech recognition applications has its own set of problems (Krause, 2000). One such problem is that over-the-phone speech applications are more difficult to translate to other languages than web applications or traditional desktop graphic user interface (GUI) applications (Head, 2000). Despite the problems, internationalizing speech applications brings with it many benefits. Internationalization of an application helps to reveal some of the peculiarities of a language, such as differences in dialects, and to shed light on the voice user interface design process (Scholz, 2000; Yan, 2000).

A long-awaited application of speech recognition, the automatic transcription of free form dictation from professionals such as doctors and lawyers, lags behind other commercial applications (Stromberg, 2000). Because of major developments in the Internet, speech recognition, bandwidth and wireless technology, this situation is changing (Bourgeois, 2000; Pan, 2000).

Speech technology has been applied to Medicine, particularly emergency medical care that depends on how quickly and accurately military or civilian field Medics access patients' background information (Kundupoglu, 2000). The United States Defense Advance Research Projects Agency organized the Trauma Care Information Management System (TCIMS) Consortium to develop a prototype system for improving the timeliness, accuracy and completeness of medical documentation. One outcome of TCIMS was the adoption of a speech-audio user interface for the prototype (Holtzman, 2000).

Applications of speech recognition technology are also being developed to improve access to higher education for persons with disabilities (Leitch and Bain, 2000). The main objectives of this project are to develop and evaluate a model for using automated speech recognition in the university classroom and to focus global attention on this concept as a method of improving access for persons with disabilities.

The seamless integration of voice recognition technologies creates a man-machine interface that has been applied to consumer electronics, Internet appliances, telephones, automobiles, interactive toys, industrial, medical, and home electronics and appliances (Soule, 2000). Of these applications, some of the most successful have been telephone-based, such as speech control in phone handsets. Continuous speech recognition has been used to improve customer satisfaction and the quality of service on telephony systems (Charry, Pimentel and Camargo, 2000; Goodliffe, 2000; Rolandi, 2000). Name-based dialing has become more ubiquitous, with phone control answer, hang-up and call management (Gaddy). These applications use intuitive human communication techniques to interact with electronic devices and systems (Shepard, 2000).

A promising application of speech recognition is to change the way the world communicates with the Web, by expanding E-commerce into V-commerce utilizing mobile interactive voice technologies (McGlashan, 2000; Pearce, 2000; Gaddy, 2000). As an enabler of talking characters in the digital world, speech recognition promises many opportunities for rich media applications and communications with the Internet (Zapata, 2000). Amid growing interest in voice access to the Internet, a new Voice extensible Markup Language (Voice XML) has surfaced as an interface for providing web hosting services (Karam and Ramming, 2000). Voice XML promises not only to speed development and expand markets for web-based services and speech recognition and synthesis services, but to spawn a new industry of "voice hosting." This business model will allow developers to build new telephone-based services rapidly (Thompson and Hibel, 2000). The voice hosting service provider will lease telephone lines to the client and voice enable a specific URL, programmed in Voice XML by the client. This model will make it possible to build speech and telephony services for a fraction of the time and cost of traditional methods (Larson, 2000).

Case Description

The NVID prototype was developed as a portable computer that employs voice interactive technology to automate and improve the environmental surveillance data collection and reporting process. Shipboard medical department personnel regularly conduct comprehensive surveys to ensure the health and safety of the ship's crew. Currently, surveillance data is collected and stored via manual data entry, a time-consuming process that involves typing handwritten survey findings into a word processor to produce a completed document. This prototype system is a compact, mobile computing device including voice interactive technology, stylus screen input capability, and an indoor readable display that will enable shipboard medical personnel to complete environmental health survey checklists, view reference materials related to these checklists, manage tasks, and generate reports using the collected data. The system operates using Microsoft Windows, an operating environment that satisfies the requirement of the IT-21 Standard to which Navy ships must conform. The major software components include initialization of the NVID software application, application processing, database management, speech recognition, handwriting recognition, and text-to-

speech capabilities. The power source for this portable unit accommodates both DC (battery) and AC (line) power options and includes the ability to recharge or swap batteries to extend the system's operational time.

The success of the NVID prototype will shed light on possible uses of speech recognition technology by the U.S. Navy for applications other than environmental surveillance.

For example, a computer system has been developed at the Naval Health Research Center to collect and document data required by U.S. Navy Medicine and the Shipboard Occupational Safety and Health Program. This system could be enhanced with the addition of voice interactive technology, facilitating the investigation of higher than expected incidences of illness and/or injuries, and to aid in preventive medicine follow-up. Follow-up support would be provided in the form of a tickler system to prompt inquiries and status checks. Voice interactive features would support the user by identifying tasks for completion and documenting the outcome of those tasks. The Navy has developed a Web-based system, the Force Health Protection System (FHPS), currently comprised of the Medical Work Station (MeWS) and the Medical Decision Support System (MDSS), for remotely or locally accessing medical data. NHRC proposes to leverage this comprehensive system to develop a Web-based repository for SAMS data. This data will be available to local shipboard personnel, Type Commanders, Environmental Health Officers (EHOs), and medical planners. Currently, medical personnel manually enter data documenting shipboard patient encounters into the SAMS system. To help automate this process, voice input could be incorporated into selected SAMS modules and remote data access could be provided. In addition, a voice-enhanced CBT (Computer-Based Training) module could be developed in selected medical domains such as preventive medicine and hazardous material exposure response.

Summary

This underway period proved to be valuable and demonstrated important lessons relevant to the success of the NVID I and II. For example, it was learned that while OPNAV Instructions set forth general, and sometimes specific, details for surveillance programs, shipboard preventive medicine technicians expand on the requirement depending on the size of the ship and the mission. Thus, it is a good idea to review shipboard procedures prior to locking down any software application for a ship. The best approach, it seems, is to design software that allows the user to modify the application to be more platform specific. Another lesson learned was that the NVID I and II must be made more intelligent in order to be accepted by the fleet. For example, the system must be made capable of handling multiple open and incomplete encounters at the same time, much like Windows allows multiple windows to be open at the same time. If lab work or x-rays exams are recorded in the plan section of the SOAP note, the system should automatically generate a request without further prompting. Other necessary improvements include the capability to access ICD-9 and SAMS. If NVID II is to be accepted Navy-wide, it must be able to link to SAMS, providing for the input and extraction of data from the SAMS database. Despite the need for improvements such as linking NVID II to SAMS, the NVID demonstrations signaled considerable successes in the development of the NVID I and II prototypes. Since NVID II automates routine procedures and generates reports and requests, medical care providers are able to use it to save time during medical encounters. The quality of patient education during an encounter is improved because the provider dictates important information in the patient's presence instead of writing it up after the patient leaves. Another success was evident in the performance of NVID I during shipboard inspections. Though the environmental surveillance functions of NVID I were still under development, the food safety module received a very good review from HM2 D, a PMT. After conducting food safety inspection in two galleys, HM2 D said that having the references onboard was beneficial and the system would save a great deal of time and improve the quality of inspections. Successes like those arising from medical encounters and food safety surveillance overshadowed the few problems impacting the NVID I and II demonstrations. One such problem was the testing of the throat microphone. Due to time limitations, this important component was not tested in a real time application. Another problem was the inability to correct programming problems onboard ship. The loss of key programming personnel led to this situation, as well as forcing a reduction in the number of prototype units available for review. Despite these few problems, the utility of NVID I and II was thoroughly demonstrated during the field test onboard the *USS CONSTELLATION*. The success of this test will inform and inspire the additional work that is still required for the pending test and evaluation.

References

Available upon request.