Appendix C

Remote Voice Training:
A case Study on Space Shuttle Applications

by

Cindy Mollakarimi & Tamin Hamid

Remote Voice Training:
A Case Study on Space Shuttle Applications

Cindy L. Mollakarimi
Lockheed Space Operations Company
1100 Lockheed Way M/S LSO-439
Titusville, Florida   32780

Tamim S. Hamid
Lockheed Space Operations Company
1100 Lockheed Way M/S LSO-439
Titusville, Florida   32780

REMOTE VOICE TRAINING:
A CASE STUDY ON SPACE SHUTTLE APPLICATIONS

1. ABSTRACT

The Space Systems Integration and Operations Research Applications (SIORA) Program was initiated in late 1986 as a cooperative applications research effort between Stanford University, NASA Kennedy Space Center (KSC), and Lockheed Space Operations Company (LSOC). One of the major initial SIORA tasks was the Tile Automation System (TAS). This system includes applications of automation and robotics technology to all aspects of the Shuttle tile processing and inspection system. This effort has adopted a dynamic engineering approach consisting of an integrated set of rapid prototyping testbeds in which a government/university/industry team of users, technologists, and engineers test and evaluate new concepts and technologies within the operational world of the Shuttle. These integrated testbeds include speech recognition and synthesis, laser imaging systems, distributed Ada programming environments, distributed relational database architectures, distributed computer network architectures, multi-media workbenches, and human factors considerations.[1]

This paper will investigate the lessons learned in remote voice training in the Tile Automation System. The user is prompted over a headset by synthesized speech for the training sequences. The voice recognition units and the voice output units are remote from the user and are connected by Ethernet to the main computer system. A supervisory channel is used to monitor the training sequences. Discussions will include the training approaches as well as the human factors problems and solutions for this system utilizing remote training techniques.

2. INTRODUCTION

An initial primary design objective for the Thermal Protection System (TPS) of the Shuttle was centered on providing a barrier to the intense thermal environment present during reentry. This objective has been fully realized with the present Shuttle tile system. During the design phase little consideration was given to optimizing the TPS design for operational maintenance efficiency. This has resulted in a TPS whose maintenance program can be characterized as being man-power intensive and time consuming. This is due to the fact that the TPS maintenance program uses manual techniques for inspection and measurement, mostly paper databases, no networking between pertinent electronic databases, manual scheduling of operational flows and a quality control and reliability program based on a paper information system.

Introducing new technologies and operational concepts into a critical system, like the Shuttle TPS, requires a careful assessment of the appropriate systems engineering approach. The SIORA Program chose a non-linear systems engineering methodology which emphasizes a team approach (design engineers, system users, technologists) for defining, developing and evaluating new concepts and technologies for the operational system. This is accomplished by utilizing rapid prototyping testbeds whereby the concepts and technologies can be iteratively tested and evaluated by the team. In addition to the skill mix of the team, it is also equally represented by the government, industry and university sectors. This later feature of the SIORA teaming is significant particularly in the areas of rapid acquisition and introduction of state of the art technologies. It also assures that the system derived from this process will be commercially viable and maintained in the future.

In considering the application of automation and robotics to the TPS several important questions must be asked. First, what technology can be applied which will produce significant productivity gains and second, what functional processes and procedures are present which lose their purpose in an automated system? The first question was surprisingly easy to address since all of the technologies were commercially available. We found that the difficult task was the integration of the technologies into an efficient and productive operational system. The first step in identifying applicable technologies was to divide the TPS maintenance system into functional process areas. This produced the following primary areas: multi-media (speech, graphics, imaging systems, test) information capture, distributed computer networks, distributed database architectures, windowed displays, software environments, simulation environment for training, and human factors considerations in system designs. The initial prototype included technologies which addressed each of the above functional areas. It was also determined that a number of functional processes would be eliminated in an automated system. These revolved primarily around procedures to validate and verify information which resided on paper databases. The interactive electronic system eliminates the need for these activities.

3. APPLICATION

The Tile Automation System consists of three major sections. First, the TPS quality control technician inspects the thermal protection system after each flight using voice data entry to identify anomalies. The inspector voices in the part number, the dimensions of the anomaly, and other necessary data which then produces an automated problem report in the central database. Second, the problem report is dispositioned by the TPS engineer using keyboard entry to identify the proper repair procedures for the particular anomaly. The problem report then proceeds through an electronic signature loop until final approval. Third, the TPS technician uses voice data entry to buy-off each work instruction and enter work control data. On specific work instructions, the TPS technician will also use automated instrumentation such as laser sensors to scan the tiles for critical dimensions of step and gap measurements between adjacent tiles. 2

The programming language environment selected by SIORA is Ada. Because this system will be in operation throughout the Space Station era, migrating to an Ada software environment is a prudent and necessary step since Space Station core systems require full utilization of Ada. An Ada environment provides excellent portability, a rich set of programming functions and tools, and a uniformity of code documentation. Also, Ada allows for multi-tasking which is critical for real-time processing. The prototype database management system chosen for this task is a commercial relational database (RELATE/DB - Computer Resources, Inc.) written in Ada. This database is also easily transportable with less than 5% equipment specific code.

SIORA has chosen a distributed hardware architecture concept. A central node will house the main database with other remote nodes on the network. The remote nodes can download the portion of the database necessary for the task at hand. Using this method, the technicians can work independently of the rest of the network. This reduces network traffic and prevents complete work stoppage in case of a single node failure. The network will be configured to adhere to ISO interface standards and will evolve to an Open System Interconnect (OSI) configuration as these standards are established. This will allow easy access to other networks when needed. The network will be connected to the NASA Program Support Communications Network (PSCN) to enable critical data to flow between essential NASA centers and Shuttle contractors.

An expert system is being developed to handle automated scheduling and quality assurance/reliability trend analysis which is critical at Kennedy Space Center. The development of the expert system will take place simultaneously with the prototyping effort such that the knowledge base can be derived from the appropriate domain experts (tile processing personnel). The implementation of the expert system will occur in the second phase of the program after the initial prototype has been fully evaluated and specified.

4. TRAINING METHODOLOGY

Dynamic engineering should be incorporated in all voice application projects if possible. This allows the users to be an integral part of the design and development of the system. Iterations can be easily incorporated and tested therefore evolving a system that is more acceptable to the users as well as being a better design.

The Voice Data Entry System (VDES) at KSC uses speech synthesis as the prompt and voice entry as the response. When executing a training pass, the speech synthesizer prompts the technician to say an utterance and the technician responds using voice data entry. The rejection/acceptance is controlled by a training supervisor who monitors the training pass by listening to the trainee over a headset. The technicians are located in their work

environment whereas the training supervisor and the voice equipment are located in another room. Therefore, the technician has no access to a terminal for the training sequences. The headset is hardwired to the voice equipment.

The International Voice Products VocaLink 4500 recognizer requires 3 passes of voice training. The vocabulary in this application consists of 96 words and takes about 2 hours per pass. The technicians were trained in a quiet environment in the first pass and subsequent passes in the working environment which is the high-bay of the Orbiter Processing Facility (OPF). There are two sets of technicians trained on the voice data entry system: first and second shift. Emphasis was placed on the second shift personel due to less interference with operations (which is probably true in many VDES applications). Also, there is lower ambient noise during second shift.

5. LESSONS LEARNED

Training time needs to be quality time. First passes were found to be critical. The quiet environment training proved to be invaluable. As the technician advanced to the working environment, the noise-cancelling microphone became a crucial piece of equipment. The ambient noise in a shuttle processing environment has considerable noise spike levels. During training, the paging system in the OPF caused recognition errors since paging is in the same frequency range as the voice input. It was important to have no paging during training and also not to test recognition while the paging system was in use. Irregular background noise is more disruptive to recognition than constant noise.

Proper inflection should be emphasized during training for the vocabulary. For example, the utterance Enter Number should not stand alone because the inflection drops off after Number. For better recognition, this utterance should be trained with digits following it (i.e. Enter Number 98432). This can be accomplished by structuring the grammar in a very tight manner and not allowing any structure that is not applicable to be incorporated into the grammar.

During remote training, speech impediments will become noticeable to the trainer. If a user pronounced a particular word inappropriately, it is difficult for the trainer to determine whether the user heard the speech synthesizer incorrectly or the user is unable to correctly pronounce the word. Also, since the speech synthesizer prompts the user in a monotone voice, some users altered their method of speaking in order to conform with the speech synthesizer's lack of inflection. When the system is in operational use, the user speaks naturally. Therefore, some problems were incurred with recognition due to lack of consistentcy during training. Words like Enter and Zero are good examples of fluctuation in speech patterns.

When performing regular training or remote training a script of the utterances is recommended in order for the user to concentrate more on the utterances rather that relying on the trainer for assistance. This familiarizes the trainee to input data from text to speech, whereas speech to speech sometimes results in mimicking of the synthesized voice and no thought is given to their own speech pattern. Another reason to use a script during remote training is that the user has to remember what the synthesizer said and repeat it. This becomes difficult for the user if the utterance is more that four words. This value fluctuates depending on the type of utterances. For example, the user usually had no trouble entering Enter Number One Two Four. This utterance is very natural to the ear and can be repeated easily. On the other hand, the user usually had difficulty repeating the utterance Enter Alpha Zebra Juliet Foxtrot. This utterance is unnatural to the ear and the users tended to either leave out one or more of the words or mix up the order of the words. This obviously led to serious problems without a script.

Users with bad recognition due to colds or variance in some utterances have the ability to perform a retrain session in real time. This results is good recognition and more importantly, the users have the control to retrain words they feel gives them trouble. This involvement in retraining allows the users to critique themselves and continuously sustain an acceptable recognition level.

Unusual synthesized voices appear to have a better effect on training than with normal voices or pleasant voices. Trainees tend to mimic the voices and do not truly talk in their natural voice pattern. With the unusual voices like DECtalk's (Digital Equipment Corporation) Ursula, Dennis, Wendy and Brat, trainees will tend to speak naturally because they find it difficult to mimic.

Training sessions for users should not be lengthy or monotonous. User motivation drops considerably as training sessions are prolonged. Designers and application engineers should consider training time in choosing the vocabulary for the application and design the grammar structure to contain a minimum number of samples or models of each utterance. Training sessions should be limited to a time agreeable to the user and the trainer. Users should be alert when performing training. If this is not achieved, operational use of the VDES will be greatly affected. Due to stress and other attributes, the voice patterns can change. A relaxed user during training will not be always relaxed during working conditions and vice versa. Users of the voice recognition should be made aware of their concentration level when inputting data. Basically, the user should be aware of the grammar structure and their own speaking volume, pitch, and pattern. Also, similar problems found in regular voice training where found in remote training. For example, poor training can result in insertion errors, nonrecognition, and misrecognition.

The user should be very familiar and comfortable with the software. To accomplish this, the users were trained on the software in front of a terminal so they could see the interaction between their voice and the software. This was found to be very critical because the user does not have access to a terminal in the working environment. Users that did not have this training had a very difficult time conceptualizing this interaction. Any indication of uncertainty of the software will be seen in the performance in recognition.

There is always the problem of resistance to change. A strategic move to convince management to convert to a VDE system would be to give them a demonstration and introduce them to the new technology. Interruption of operational work for training on a new system and sacrificing key personnel where their absence results in loss of work is definitely not tacticly smart. A training strategy is essential in converting existing methods to the new technology. A training module for the Tile Automation System is being established with the University of Central Florida from an Industrial Engineering point of view. This training module will be designed for upper management all the way down to the technicians. The transition to a VDE system is critical in the Thermal Protection System of the shuttle because of the small amount of trained

technicians. The amount of time to train a technician for the VDE system has to be carefully scheduled because the absence of a technician could hamper scheduling requirements.

The headphones used on this project are Shure noise-cancelling microphones. These headphones have an amplifier built in to boost the signal due to the long transmission distance. The problems found associated with the headphones are as follows:

1) Inconsistency of microphone positioning where it affects recognition

2) The headphones break rapidly when they are constantly used in the working environment.

3) The headphone's design does not subtract the ambient noise effectively. Continuous paging in the high-bay during training passes or operational use is a problem, unless the training first pass is performed well.

One solution that is being investigated to the noise problem in the working environment is an ear microphone. This earphone fits snugly into the external auditory canal and drives acoustical energy through the Eustachian tube. Foam is used to cancel out the ambient noise and also to keep the earphone intact. A small case would be attached to the user, this case contains the processor and battery. The testing of the earphone was conducted with the International Voice Products Series 4000 recognizer. Test results indicated that the earphone was highly sensitive, even after changing gain levels to different settings. The earphone had constant recognition errors with coughs, clearing of the throat, and normal conversations. This technology should be closely watched since it could greatly improve recognition and its human factors aspect of it.

6. SUMMARY

The non-linear system engineering methodology, with its team approach and rapid prototyping techniques, has clear advantages for the design of large complex systems as well as for the upgrading and evolution of existing systems. The SIORA Program will thoroughly test the methodology on an existing system, the Shuttle processing at KSC, while the rapid prototyping efforts for a number of aspects of Space Station Program will test the effectiveness of the methodology on a new, complex system. The future space program requires a new and innovative approach to system engineering such that operational systems are functionally productive and cost effective.

The Tile Automation System has tested remote training for voice data entry and has found that it can be successful. There are other technologies still to be tested in this application such as an RF network, digitial voice transmission, and dual voice users interacting with the same application software.

7. ACKNOWLEDGEMENTS

8. REFERENCES

1. M. J. Wiskerchen, "Telescience -- A Space Station Goal", SPIE Cambridge Symposium on Optical and Electro-Optical Engineering, 1985

2. C. L. Mollakarimi and M. A. Yvanovich, "Voice Data Entry on Space Shuttle Processing", American Voice Input/Output Society Conference, 1987