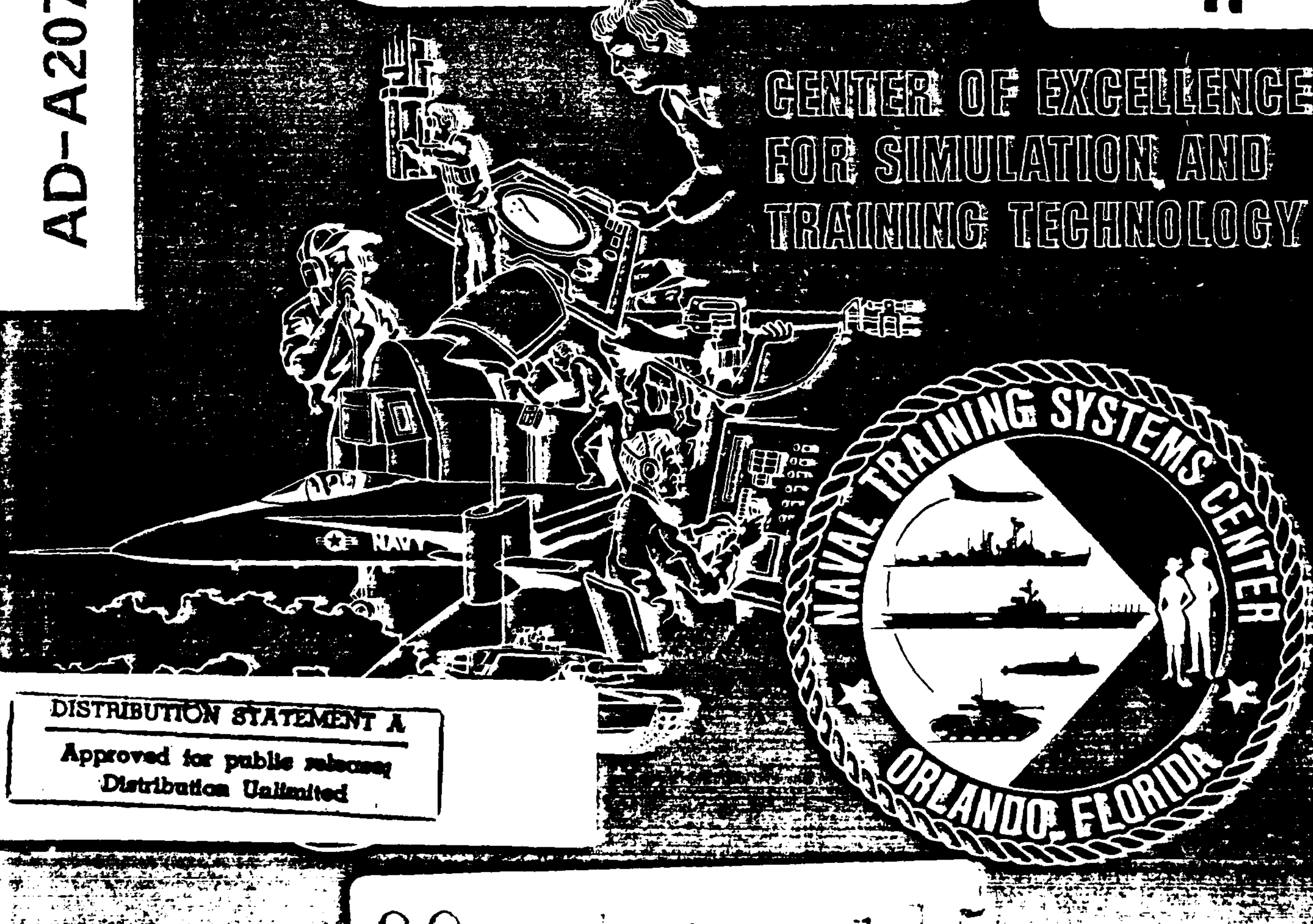
NAVAL TRAINING SYSTEMS CENTER
ORLANDO, FLORIDA

SPECIAL REPORT SR89-01

MICROCOMPUTER SYSTEM INTEGRATION FOR AIR CONTROL TRAINING

JANUARY 1989



UNCLASSIFIED

[illegible]ECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
U[illegible]assified

1b. RESTRICTIVE MARKINGS

2a. [illegible]RITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
NTSC SR89-01

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Naval Training Systems Center

6b. OFFICE SYMBOL (If applicable)
Code 712

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS (City, State, and Zip Code)
Orlando, FL 32826-3224

7b. ADDRESS (City, State, and Zip Code)

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Naval Air Systems Command

8b. OFFICE SYMBOL (If applicable)
933G

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State and Zip Code)
Washington, DC 20361

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
63733N	W1773	7785	DN7080034

11. TITLE (Include Security Classification)
M[illegible]rocomputer System Integration for Air Control Training (U)

12. [illegible]SONAL AUTHOR(S)
Hamel, Cheryl J., Kotick, D., and Layton, M.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 8701 TO 8712	890125	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	06	
23	02	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Speech recognition, voice synthesis, air intercept control, simulation, instructional techniques. (SDW)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The research was designed to demonstrate a microcomputer system configuration for air control training. The objective of the research was to systemmatically integrate off-the-shelf components of microcomputer and microprocessor technology into a model prototype system for air control training. A subgoal was to examine voice input/output technology for its potential use in a microcomputer-based air control training system. The report describes the microcomputer system PC/NTDS (Personal Computer/Naval Tactical Data System) used for Air Intercept Control (AIC) training. It describes the automated speech recognition and speech generation technology chosen to interface with the system. Data collected using several subjects who completed AIC lessons using the system are summarized. Advantages and disadvantages of using voice input/output technology for air control training are discussed, and recommendations for future research using the microcomputer testbed are provided. Keywords:

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Cheryl J. Hamel	407-380-4825	Code 712

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

SPECIAL REPORT SR89-01

MICROCOMPUTER SYSTEM INTEGRATION FOR AIR CONTROL TRAINING

JANUARY 1989

Cheryl J. Hamel
David Kotick
Mark Layton

NAVAL TRAINING SYSTEMS CENTER
Orlando, Fl 32826-3224

W.A. RIZZO, Head
Human Factors Division

H.C. OKRASKI, Director
Research and Development Department

Special Report 89-01

EXECUTIVE SUMMARY

INTRODUCTION

The research was designed to demonstrate a microcomputer system configuration for air control training. The objective of the research was to systematically integrate current off-the-shelf components of microcomputer and microprocessor technology into a model prototype system for air control training. A subgoal was to examine voice input/output technology for its potential use in a microcomputer based air control training system.

METHOD

The microcomputer system PC/NTDS (Personal Computer/Naval Tactical Data System), an experimental system for air intercept control (AIC) training, was interfaced with automated voice recognition and speech generation technology. Recognition accuracy data was collected from five subjects who completed AIC lessons.

Besides the addition of voice technology, PC/NTDS was configured to include a pseudopilot station which consisted of a monochrome monitor and standard PC keyboard. The communication interface with the student controller station was a two-way radio headset.

RESULTS AND DISCUSSION

The product of the research was a proof-of-concept testbed for the application of microcomputer technology to air control training. The testbed demonstrates a student controller station and an instructor/pseudopilot station. Optional voice equipment at the student controller station may be substituted for the pseudopilot. For the five subjects who completed AIC component skill lessons using voice input, errors of misrecognition averaged under five percent in all lessons. Errors of nonrecognition, where input was not accepted and had to be repeated, were relatively high, ranging from 4 to 17 percent.

Several system characteristics were found to be important to successful implementation of voice input/output technology: the user interface during enrollment and system training, sensitivity to background noise, and speaking habits of the user.

The advantage of using voice input/output (I/O) at the student controller station is that it allows the practice of communications without the need for a pseudopilot. Disadvantages of voice I/O are centered around the importance of human performance to obtain system reliability. Several variables, including preparation to use voice I/O and proficiency on a training task, were found to be directly related to the utility and reliability of voice technology for air control training.

Recommended directions for research using the microcomputer system testbed focus on making the testbed more generic to enable research on other air control training applications and other applications involving voice technology. Specific recommendations are also included to improve the efficiency of automated voice recognition systems used with air control training devices.

ACKNOWLEDGEMENTS

Appreciation is expressed to Anna Fallon for her support in the testing of experimental subjects on the voice recognition systems.

Special Report 89-01

TABLE OF CONTENTS

SECTION		PAGE
I	INTRODUCTION	1
	Objective	1
	Background	1
	Voice Input/Output Technology	2
	Organization of this Report	3
II	METHOD	4
	Description of PC/NTDS for Air Control Training	4
	Selection of Voice Input/Output Technology	4
	Air Intercept Control Vocabulary	5
	Voice Enrollment Testing Procedures	5
	Voice Output Testing Procedures	7
	Voice Application Testing Procedures	7
	Pseudopilot Testing Procedures	9
III	SYSTEM HARDWARE CONFIGURATION	11
	General Overview	11
	Controller-Pseudopilot Configuration	11
	Voice-Operated Pseudopilot Configuration	13
IV	RESULTS	15
	Voice Enrollment	15
	Recognition Errors During Skill Training	16
	Intersubject Variability	18
	Voice Output Results	21
	Advanced Air Intercept Control Training	21
V	DISCUSSION	22
	Microcomputer Integration	22
	Training Alternatives	23
	Voice Input/Output Technology Application	24
VI	RECOMMENDATIONS	26
	Future Research Using the Testbed	26
	Use of Voice Input/Output Technology with Training Devices	26
	REFERENCES	29
	APPENDIX A. Interstate Vocabulary List and Syntax	A-1
	APPENDIX B. Schematic Drawing of the Controller Interface Box Circuit Card	B-1

LIST OF ILLUSTRATIONS

FIGURE		PAGE
1	Pseudopilot configuration	12
2	Voice-operated controller station	14
A-1	Interstate syntax structure for the air intercept control vocabulary	A-2
B-1	Schematic drawing of the controller interface box circuit card	B-1

LIST OF TABLES

TABLE		PAGE
1	Features of Votan and Interstate Continuous Voice Recognizers	6
2	Percent Nonrecognitions and Misrecognitions forThree Air Intercept Control Lessons	16
3	Total Percent Recognition Errors and Response Time (RT) in Each Air Intercept Control Lesson	17
4	Performance and Recognition Accuracy (Subject x Lesson)	19

SECTION I

INTRODUCTION

In Air Intercept Control (AIC) and Air Traffic Control (ATC) training, time and equipment deficiencies limit drill and practice of skills in the schools. Continued training is expected to occur on the job. For ATCs, on-the-job training may occur months after graduation from ATC school. For AICs, training exercises aboard ship are often few and far between. A training technology which can increase training efficiency in the schools and facilitate retention of basic air control skills over periods of non-use is required for fleet readiness.

This 6.3 advanced development effort was designed to demonstrate a microcomputer system configuration for air control training. A 6.2 exploratory research project conducted by the Naval Training Systems Center (NAVTRASYSCEN), developed an experimental training system which demonstrated the feasibility of microcomputer simulation of a radar task. This project builds on the capabilities demonstrated in the 6.2 development effort and researches new technologies to expand the experimental training system.

OBJECTIVE

The objective of the research was to systematically integrate current off-the-shelf components of microcomputer and microprocessor technology into an advanced development model prototype system for air control training. A subgoal was to examine automated voice recognition and speech generation technology for its potential use in a microcomputer- based air control training system.

BACKGROUND

Research on perceptual skills training suggests that theory-based drill and practice techniques employing microcomputers can produce automatic skills which are resistant to stress and forgetting (Schneider, 1982). NAVTRASYSCEN research is developing drill and practice methods to train spatial visualization for determining intercept heading calls for AIC training. The methods employ component skill training and automated performance feedback on an IBM PC/XT microcomputer. The system can compress component training through problem time compression, allowing the trainee to experience over ten times as many trials at component tasks per unit time as real-time training. The component skill training modules for training spatial visualization skills in the AIC task uses a simulated radar plot with aircraft tracks and Naval Tactical Data System (NTDS) symbology. The screen also displays instructions and feedback for trainees. In addition to the component skill training modules, the program simulates the whole AIC intercept task. A trainee can control both a friendly and enemy aircraft and perform a nearest collision intercept.

The original configuration of the microcomputer system described above requires the trainee to input responses which activate aircraft via a standard PC keyboard. This is an acceptable form of response input for that part of the training strategy which involves massive drill and practice of visual components of the task before practicing the whole task. However, to adequately train the whole task, the next step in the continued development of the training system must be the inclusion of a means to practice controller-pilot communication.

The AIC school currently conducts training by stimulating operational equipment which requires operators at pseudopilot stations to activate aircraft via keyboard inputs and simulate pilot communication. To advance development of the experimental system, it was hypothesized that one or more microcomputers could be configured to produce a trainee-pseudopilot arrangement similar to that used in the AIC schools. It was also hypothesized that a second system could be configured for comparison which interfaced a voice input/output unit with a single microcomputer. The voice system could replace keyboard input during component training and pseudopilot functions during whole task training.

VOICE INPUT/OUTPUT TECHNOLOGY

Voice recognition systems have a variety of characteristics which must be considered for each application. First, recognition systems are speaker-independent or speaker-dependent. Speaker-dependent systems require speakers to train the system to their own voice. The vocabulary is repeated until an acceptable voice template is formed. The template is used by the system to match utterances to the template to obtain recognition. Speaker-independent systems do not require speakers to enroll their own voice. Instead, recognition is based on a voice template obtained from sampling a large variety of speakers. The technology of speaker-independent systems has not advanced to the point of making them a viable alternative. Currently, vocabularies on such systems are limited to digits and a few words (Christ, 1984).

Second, voice recognition systems recognize continuous speech or discrete words. Discrete recognizers require a pause between each word entered, creating an unnatural manner of speaking. Discrete recognizers can handle large vocabularies and show higher accuracy than continuous recognizers.

There are two types of speech generation systems: digitized speech and synthesized speech. Digitized speech systems play back audio recordings. Synthesized speech systems change text to speech by electronically reproducing the phonemes which make up human speech. Digitized speech requires more storage space than synthesized speech, but it is more easily recognized and sounds more human (Christ, 1984).

Voice recognition and speech synthesis are attractive technologies to be investigated for their potential application to interactive training systems such as simulated air control training. Controller-pilot communication can be simulated by a combination of voice input and output and the relatively small, standardized air control vocabulary fits within the restrictions of most off-the-shelf voice recognition systems. Past attempts to develop voice-controlled air control training systems have met with some success. An experimental prototype training system for precision approach radar controllers used a discrete voice recognizer to successfully demonstrate task simulation, interactive instruction, automated performance measurement and adaptive syllabus control, using a fairly small vocabulary (McCauley, 1984). Another experimental prototype training system, developed under the sponsorship of the NAVTRASYSCEN, trained air intercept controllers. This system used a continuous voice recognizer interfaced with a minicomputer and performed similar instructional functions (Halley, 1981). Both systems encountered a number of problems during operation; nevertheless, they demonstrated the potential of integrating voice recognition and synthesis and interactive instruction.

With the anticipation that voice technology had advanced significantly in the last five years, an attempt was made in this research to investigate existing, off-the-shelf voice systems for their potential application to air control training using a microcomputer-based system.

ORGANIZATION OF THIS REPORT

In addition to this introduction, the report contains five sections and two appendices. Section II describes the methods used to select equipment and describes the application of the micromputer to AIC training. Section III describes the final equipment configuration which made up the testbed. Section IV reports the results of voice recognition tests of the system. Sections V and VI provide discussion and recommendations on the potential uses of the system for air control training. Appendix A contains the AIC vocabulary list and syntax structure used with the voice recognition system and Appendix B is a schematic drawing of the circuit card in the simulated controller console.

SECTION II

METHOD

This section describes 1) The microcomputer system PC/NTDS (Personal Computer/Naval Tactical Data System) used for Air Intercept Control (AIC) training, 2) The procedures used to select an automated voice input/output system to interface with the system and the procedures used to configure the two-station (controller and pseudopilot) AIC training system, and 3) the procedures used to evaluate the microcomputer configuration's reliability and utility for AIC training.

DESCRIPTION OF PC/NTDS FOR AIR CONTROL TRAINING

PC/NTDS is a state-of-the-art simulation of the air intercept control task, operating on an IBM personal computer. The simulation functionally duplicates several Naval Tactical Data System (NTDS) displays and controls used in training and combat. The system includes an IBM PC/XT computer, trackball, simulation programs, and printer. The PC/NTDS system provides many features to speed skill acquisition. Training features include: a) introductory frames explaining procedures and objectives, b) windows that display information and feedback during a simulation, c) menu systems to move between lessons, d) automated performance measurement, e) simulation control to speed up or slow down simulated time, f) instant replays of segments, and g) hard copy printouts of tactical displays.

SELECTION OF A VOICE INPUT/OUTPUT TECHNOLOGY

Several voice input/output systems were viewed and analyzed before a purchase decision was made. After a review of the sales literature, two isolated voice recognition systems were selected to receive further scrutiny through company demonstrations. These systems were Kurzweil Voice System (KVS) (Kurzweil Applied Intelligence, Inc.) and VoiceScribe 1000 (Dragon Systems, Inc.). The Kurzweil system costs approximately $5500 and has a 1,000 discrete word vocabulary size. The Dragon system costs approximately $2000 and has the same vocabulary size. Both advertize recognition accuracy rates approaching 98 percent.

The discrete voice recognizers were ruled out after demonstrations revealed that discrete input would not meet AIC training requirements. Distinct pauses between utterances produce choppy phrases and sentences incompatible with the style of controller input. Although it is possible to use relatively long phrases as one utterance in a discrete recognition system, enrollment time, during which the vocabulary is trained, is significantly increased. For instance, 3-digit headings can be input as one utterance, but it is then necessary to train two passes on 360 different headings. With a continuous recognition system, only two training passes on ten digits would be required.

After considering the cost and features of several continuous voice recognition systems, two systems were selected for purchase: Vocalink 4000 (Interstate Voice Products, Inc.) and VPC 2000 Voice Card (Votan, Inc.). Table 1 lists the cost and features of each. Votan included a voice digitizer/ playback unit for voice output on the board itself and so it was included in the purchase price. Interstate required purchase of a speech synthesizer at additional cost. The synthesizer was sold by Interstate and developed by Street Electronics.

AIR INTERCEPT CONTROL VOCABULARY

A vocabulary was developed that would encompass all of the words or phrases used in a selected group of AIC lessons. The lessons trained aircraft heading estimation, reciprocal of heading, intercept heading calls, and control of fighter and bogey in a two-plane nearest collision intercept. The vocabulary consisted of words and phrases and the digits 0 through 9. "Heading right 010 for the bogey," is a sample phrase from the vocabulary. The complete vocabulary list and syntax developed for the Interstate system are shown in Appendix A. Syntax development reduces the number of vocabulary inputs that compete for recognition at various points in the syntax. This decreases the time required for matching inputs to templates, which reduces system response time and increases recognition accuracy.

The Votan system incorporated the same vocabulary as Interstate, but the capability to program syntax was not utilized because it required use of the Voice Programming Language (VPL). To use this language requires a minimal knowledge of computer programming. Thus, with the use of the Votan, every word in the vocabulary competed with every other word during recognition.

VOICE ENROLLMENT TESTING PROCEDURES

Voice enrollment procedures are a significant feature of voice recognition systems when the application involves large numbers of trainees who will be using the system. A unique voice template is required for each trainee in the course, and so the time used to create a template takes away from time devoted to the curriculum. The ease with which the enrollment is accomplished facilitates trainees' use of the system and reduces the enrollment time.

Four subjects, different from the five subjects used later for voice testing, were used to study voice enrollment procedures using the two systems. One male and one female had their voices enrolled on either the Interstate or Votan system. An experimenter assisted the subjects during enrollment procedures.

Table 1

Features of Votan and Interstate
Continuous Voice Recognizers

	INTERSTATE	VOTAN
Recommended number of training samples per word:	2	2
Amplitude control (input)	no	yes
Manual control of sequencing of training?	no	yes
Extraction of templates:	automatic	automatic/manual
Procedure for adding/ changing one word?	yes	yes
Automatic refinement of template during application?	no	no
Speaker-independent recognition?	no	no
Recognizes continuous speech?	yes	yes
Maximum size of vocabulary:	100 words	64 words per set
Programming of syntax?	yes	yes, with Voice Programming Language
Compatible with higher-order languages?	yes	yes
Set acceptance threshold:	automatic	manual
Description:	17x4x12 inch stand-alone	IBM long slot board
Intended computer?	computer with RS232C interface	IBM PC or compatible with 256K memory
Voice output included?	no	record/playback
Power consumption:	45W	9W
Base cost:	$5200	$2000

After the enrollment period, a short practice recognition test, programmed by the manufacturer, was administered.

VOICE OUTPUT TESTING PROCEDURES

The Interstate/Street Electronics speech synthesizer was low cost compared to other synthesizers on the market. The lower quality of the synthesizer was not expected to be detrimental because the pilot vocabulary for AIC training is extremely limited. Text to speech algorithms produced the required "pseudopilot" speech. Preliminary work with the synthesizer, however, revealed major problems. The words sounded very mechanical and multiple word phrases sounded "run together." The experimenters made some attempts at phonetic speech output, i.e., words were typed in as they sound rather than as they are spelled, but this method was time-consuming and was not producing acceptable results. So the synthesizer was dropped from the experiment.

The speech digitizer/playback unit, received as part of the Votan card, was used to record the voice of one of the experimenters. Digitization rate is selectable on the Votan. A mid-range digitization rate was used after some experimentation with the board.

A small vocabulary was developed to test the performance of the Interstate/ Street Electronics and Votan voice output systems. The vocabulary consisted of words or phrases such as "Roger", "Missile enabled", and "Missile away." Evaluation software was written to test both voice output systems.

VOICE APPLICATION TESTING PROCEDURES

The Interstate Vocalink 4000 input system was interfaced with the AIC training software, and system performance during AIC training conditions was evaluated. Subjects received training to criteria on several visual components of the AIC task. During each AIC training session, time of day, background noise conditions, task performance measures, and system recognition errors were recorded. Statistical tests were not performed due to the small number of subjects who participated in the research ($n = 5$). The observations were treated as case studies from which data could be extracted and used to form human factors generalizations pertinent to the application of voice input/output to training. System performance was the central area of concern. Subjects' performance on the AIC tasks was measured to observe the relationship between level of training of the subject and voice recognition accuracy of the system.

Subjects

Five students, two male and three female, ages 17-24 years, participated in the research. Their areas of study were computer programming and psychology. None had prior experience with automated voice recognition systems. A female experimenter completed voice enrollment and AIC training to become familiar with the procedures and then tested the five subjects over a period of a month.

Voice Enrollment

Sessions averaged 25 minutes, during which time subjects were given practice followed by two training passes using the AIC vocabulary. At the beginning of the session, the experimenter helped the subject adjust the headset and position the microphone a distance of three fingers away from the mouth. The subject then proceeded to read a series of words and phrases which appeared on the CRT. The subject was asked to speak in a natural manner; background noise was minimal. After five minutes of practice, during which the experimenter coached the subject on how to speak properly, the enrollment session began. During enrollment, the experimenter controlled voice entries by accepting or rejecting each entry. Entries were rejected when the subject used unacceptable vocabulary or stopped before the complete phrase was given.

Skill Training

The software consisted of three lessons designed to reinforce basic air intercept control skills. All lessons were conducted using a simulated radar screen with a graphic compass rose. Verbal instructions and a demonstration were given to each subject prior to the beginning of each lesson. Lessons consisted of a series of drill and practice trials in which subjects input a short response and received immediate feedback. Accuracy rather than speed was stressed and subjects were told they would be given as many training sessions as necessary to reach criterion. All lessons were 20 minutes long.

Accuracy of the voice recognition system was measured by recording misrecognitions and nonrecognitions. Each lesson required one input which was to be spoken as a complete phrase or sentence. If a subject said part of the input, the subject was required to repeat the input and the error was not counted as a system recognition error. Misrecognitions were defined as the system recognizing an input different from the subject's response. Nonrecognitions were defined as a rejection of an input by the system. In such instances, the system would wait until another response was entered. All errors were recorded manually by the experimenter who sat next to the subject during the session. If the system did not recognize three sequential inputs the experimenter typed in the third input. Thus, no more than two nonrecognitions were recorded on any one trial. Percent misrecognitions and percent nonrecognitions were recorded.

<u>Heading Lesson</u>. The object of this lesson was to identify the heading of an aircraft, as seen travelling on the radar screen, to the nearest 10 degrees. The response required was a phrase, "Heading xxx for the bogey," which included a mandatory 3-digit heading call. Automated feedback was provided on

response time and response accuracy. When a subject's average error in a session was approximately 5 degrees, the subject was started on the next lesson in the next session.

Reciprocal Lesson. The object of this lesson was to identify the reciprocal heading of an aircraft, as seen travelling on the radar screen, to the nearest 10 degrees. A reciprocal is 180 degrees off the heading, or in other terms, is directly opposite the heading on the compass rose. The subject also had to turn the aircraft to the reciprocal heading by indicating left or right. The response required was a phrase, "Viper/snake left/right xxx for the bogey" which included a mandatory 3-digit heading call. When a subject's average error in a session was approximately 5 degrees and/or direction calls were approximately 95% correct, the subject was started on the next lesson in the next session.

Intercept Lesson. This lesson employed skills gained from the previous lessons and was the most difficult of the component skill lessons. Two aircraft appeared on the screen, the bogey and fighter. It was explained to the subject that the bogey would remain on a fixed path and that the objective of the lesson was to turn the fighter to a heading that would intercept the bogey. The response required was a phrase, "Heading left/ right xxx for the bogey" which included a mandatory 3-digit heading call. All subjects received two intercept lessons.

Voice Recognition Feedback and Error Correction. During skill training, visual and auditory feedback were used to indicate system voice recognition. Once a phrase was said in its entirety, the entire phrase was displayed on the screen and an auditory signal was provided in the headphone. The response was automatically entered. Subjects were not able to confirm entries, e.g., by using the ENTER key, nor were they able to correct an error in the middle of a phrase. Individual units of the phrase did not appear on the screen as they were being spoken. If a subject stopped short of a completed phrase, nothing was displayed and the subject simply started the phrase over again from the beginning.

PSEUDOPILOT TESTING PROCEDURES

The instructor and pseudopilot stations were configured to represent the traditional AIC training method in which a student controller communicates to a pseudopilot by radio and the pseudopilot moves targets via a keyboard. A description of the configuration is presented in Section III.

A two-plane control scenario was developed to demonstrate use of the instructor/pseudopilot station. In this scenario, the subject controlled both aircraft using the same vocabulary developed for the component skill training. The subject first turned the bogey to a given heading by saying to the pseudo-pilot, "Snake, left/right xxx as bogey," then the subject used Function 2 on the control box to obtain the bogey's bearing and range. The use of Function 3 (Sequence) which gave readouts of bogey and fighter information was optional.

Finally, the subject gave an intercept heading call to the fighter by saying the message, "Viper, left/right xxx for the bogey." The scenario ended with two commands, "Enable missile" and "Fire missile." The capability to "blow up" the bogey was added for motivational purposes only. Commands were issued over a two-way communication system. The pseudopilot was instructed to respond "Roger" to all heading commands and to execute the required action. The pseudopilot also said "Missile enabled" in response to the "Enable missile" command and "Missile away" to the "Fire missile" command.

SECTION III

SYSTEM HARDWARE CONFIGURATION

This section describes the computer hardware and programming languages used in the development of the microcomputer testbed. First, a general overview is provided followed by separate descriptions of the controller-pseudopilot configuration shown in Figure 1 and the voice-operated pseudopilot configuration shown in Figure 2.

GENERAL OVERVIEW

The PC/NTDS microcomputer configuration is composed of one IBM PC/XT central processing unit with one 360 kilobyte floppy disk drive and one 10 megabyte fixed disk drive. With the addition of a WEN multifunction card with 384K memory, the total memory capacity of the CPU is 640K. The CPU also requires an additional 8087 math coprocessor.

The IBM monochrome monitor has a 700 x 700 pixel resolution with the use of a TECMAR Graphics Master board. The PC/NTDS perceptual skills program will not run with a standard IBM monographics card; however, an off-the-shelf graphics package can be installed to allow the program to run with almost any graphics card and monitor. Future development of the testbed should begin with this addition.

A WICO trackball is used with a TECMAR Labtender input/output board, and an Epson RX-80 dot matrix printer is used with an asynchronous serial interface board.

The PC/NTDS programs are written in the PASCAL programming language, using the IBM PASCAL compiler. The disk operating system is IBM DOS version 3.0.

CONTROLLER-PSEUDOPILOT CONFIGURATION

The controller-pseudopilot configuration is shown in Figure 1. The configuration consists of two stations, one for the student controller and one for the operator who controls targets and transmits pilot messages (pseudopilot). The hardware consists of one computer and two terminals. The software computational load does not require more CPU power than what the one computer provides, and using only one computer is less expensive.

The student controller station consists of a switchbox which emulates AIC NTDS functions and a CRT which displays the radar simulation and other instructional features. The 12-switch interface box can be programmed to emulate 12 functions. At the current time seven switches are programmed. The functions programmed are: F1 - Hook target, F2 - Get bearing and range of the bogey, F3 - Display information in aircraft sequence list, F7 - Turn fighter to the left, F8 - Turn fighter to the right, F9 -

Compress scenario time, F10 - Go to the next scenario. A schematic drawing of the controller interface box is provided in Appendix B.

The pseudopilot station consists of the IBM standard keyboard and a CRT slaved to the controller CRT. With this arrangement, the operator uses the keyboard to manually control targets and watches the same display as the controller. In many air control training devices, the pseudopilot does not have access to a visual display. This addition may have training value for the student controller who is required to play pseudopilot during course time.

The student controller and pseudopilot communicate using audio headsets equipped with noise-cancelling microphones.

VOICE-OPERATED PSEUDOPILOT CONFIGURATION

The voice-operated pseudopilot configuration is shown in Figure 2. With this arrangement there is one student controller station equipped with a voice input/output system (either the Votan or Interstate). The voice system takes the place of the pseudopilot. The voice recognition system activates targets and the speech output system is programmed to emit pilot transmissions at appropriate points in the simulation.

The two voice systems examined in the testbed were the Votan, Inc. VPC 2000 voice recognition system and the Interstate Voice Products Vocalink 4000 voice processing system. The two systems included in the testbed represent off-the-shelf low cost connected voice recognition/ output systems. The Interstate system has a separate speech output module, whereas the Votan unit incorporates voice digitization and playback on the same card as the recognition unit.

The Votan VPC 2000 unit is a single peripheral printed circuit card (IBM PC format) which is installed in a single slot on an IBM PC/XT or compatible. It communicates to the host system via the PC bus and to the outside world by audio cabling to the microphone (voice input) and speaker (voice output).

The Interstate recognizer is a stand alone unit with host connections made by a serial RS232C link. The speech output unit is also a stand alone unit connected serially via RS232C. Both recognition systems in the testbed execute the vendor-supplied keyboard emulation software.

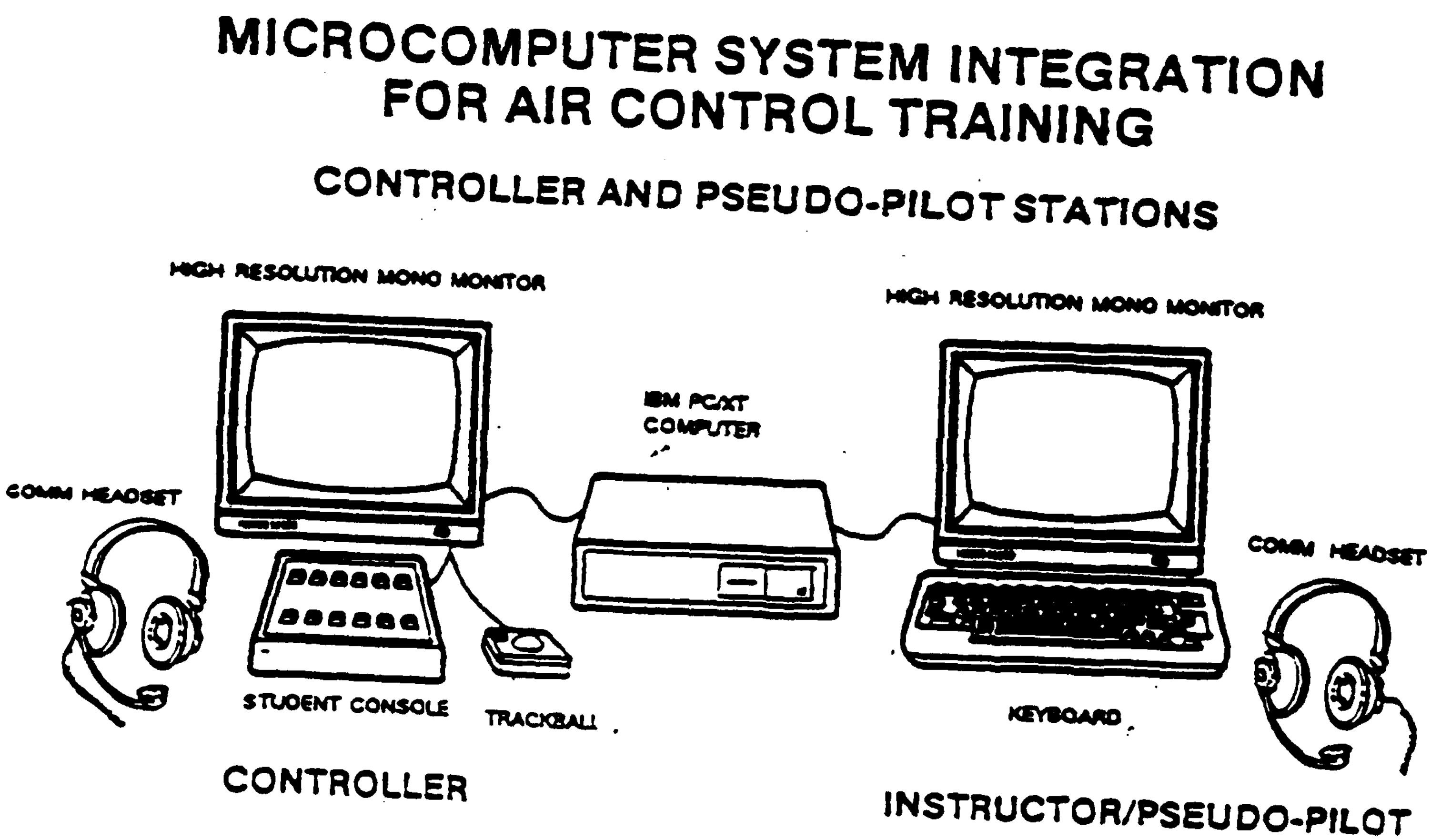


Figure 1. Pseudopilot configuration.

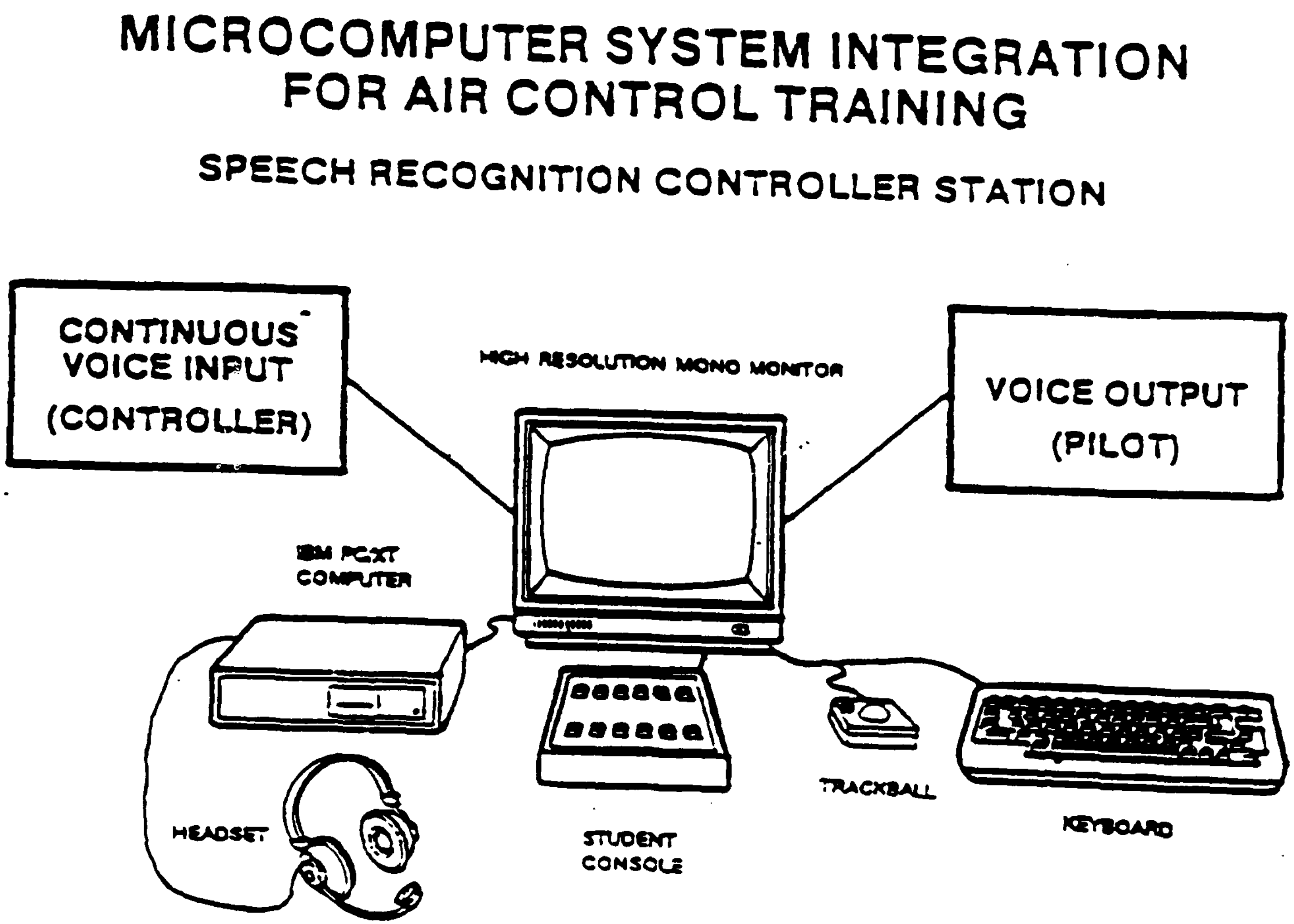


Figure 2. Voice-operated pseudopilot.

SECTION IV

RESULTS

This section describes the results of the comparison of the Votan and Interstate voice recognition systems during voice enrollment procedures which led to the selection of the Interstate system for application to air intercept control (AIC) training in the NTSC testbed. The section next describes the voice recognition results obtained from subjects who used the Interstate system to conduct AIC training.

VOICE ENROLLMENT

Voice enrollment procedures using the Votan and Interstate systems were compared. Vocabulary input, creation of syntax, and voice training procedures were examined.

Before voice enrollment began, it was necessary to input the vocabulary to be trained. Using the Interstate Vocalink 4000 system, input of the 19 words/phrases and digits to be used in the AIC application was accomplished in approximately 20 to 30 minutes. This included creating the translation table and syntax file using Interstate-provided programs and downloading the files to Interstate to make the master cartridge. Next was the voice enrollment which, for each subject, took about 25 minutes due to the automated voice enrollment procedures incorporated into the Interstate system. The system includes software which displays vocabulary units either singly or in various combinations during voice enrollment. Subjects were required to repeat the units or phrases as they were displayed and then were given the option to accept or reject their input. As the Interstate manual suggested, two passes through this automated voice enrollment procedure were given to each subject. The recognition accuracy test, given at the end of voice enrollment, showed an accuracy level of 100% for two subjects.

Voice enrollment using the Votan VPC 2000 was relatively more complicated and time-consuming because of the lack of automated voice training procedures. Vocabulary input consisted of typing in all digits, words, and phrases in the vocabulary list. Syntax was not developed due to the experimenter's unfamiliarity with the Votan Programming Language (VPL). Voice training was entirely manual, i.e., the experimenter told the subject whether to say an isolated input or a string of inputs, decided when inputs had enough repetitions, and developed strategies on how to combine inputs to obtain the best recognition. The Votan system also allowed manual control of input level and acceptance threshold. Input level accommodates various voice amplitudes. Acceptance threshold delivers the desired ratio of nonrecognitions to misrecognitions. A higher threshold results in better rejection of invalid words (less misrecognition errors) at the expense of a greater frequency of rejection of valid words (more nonrecognition errors).

Two subjects who used the Votan system and were given 45 minutes to an hour for voice enrollment, did not reach acceptable accuracy levels on the automated practice recognition test provided by the manufacturer. Part of the problem was that the experimenter did not have the skills and knowledge required to conduct voice enrollment. On the basis of this experience with the Votan system, it was not used in the research; however, the card was installed in the testbed for future work. One of the authors, an electrical engineer, spent several days experimenting with strategies and procedures for voice enrollment using Votan and was able to get acceptable recognition (> 90% accuracy) when he trained himself on the system.

RECOGNITION ERRORS DURING SKILL TRAINING

Recognition errors were determined on a message-by-message basis. Each trial counted as one message. A nonrecognition or misrecognition of a message was counted as an error for that trial. The results of five subjects given AIC training are shown in Table 2. In the heading lesson, subjects showed an average of 336.8 trials to criterion. Nonrecognitions, i.e., where an input was not acknowledged, averaged 4.6 percent of the total number of trials and the range among subjects was 2.0 to 6.5 percent. Misrecognitions averaged 3.4 percent, with a range of 1.0 to 6.5 percent. Percent recognition errors were determined by the number of trials showing an error divided by the total number of trials.

Table 2

Percent Nonrecognitions and Misrecognitions of Messages for Three Air Intercept Control Lessons

Lesson	Average Number of Trials	Percent Non recognitions	Percent Mis-recognitions
Heading	336.8 (235 - 497)	4.6 (2.0 - 6.5)	3.4 (1.0 - 6.5)
Reciprocal	168.8 (129 - 185)	17.0 (13.0 - 27.0)	4.4 (0.3 - 8.0)
Intercept	82.8 (62 - 98)	15.4 (1.0 - 31.5)	4.2 (0.0 - 9.5)

Note. Ranges are shown in parentheses.

In the reciprocal lessons, subjects showed an average of 168.8 trials to criterion. Percent nonrecognitions averaged 17.0, with a range of 13 to 27 percent. Misrecognitions averaged 4.4 percent, with a range of .3 to 8.0 percent.

In the intercept lessons, subjects showed an average of 82.8 trials to criterion. Nonrecognitions averaged 15.4 percent with a range of 1 to 31.5 percent. Misrecognitions averaged 4.2 percent with a range of 0 to 9.5 percent.

The data in Table 2 reveal that percent misrecognitions remained fairly constant from lesson to lesson. Percent nonrecognitions were higher and did not remain constant throughout training. The Interstate acceptance threshold is set high so that the system accepts only a very small range of variation of input, thereby increasing the likelihood of rejections. Observations by the experimenter during training revealed that many rejections were caused when subjects started a phrase and then paused too long before saying the direction or the heading of the aircraft. This operator error and others of this type were most frequent in the first session of each new lesson when subjects were at the low end of the learning curve.

Table 3 shows misrecognitions and nonrecognitions combined to obtain total percent recognition error over sessions.

Table 3

Total Percent Recognition Error and Response Time (RT) in Each Air Intercept Control Lesson

	SESSION NUMBER			
	ONE	TWO	THREE	FOUR
HEADING LESSON				
Average RT (Seconds)	6.0	4.6	4.8	4.6
Percent Error	12.4	6.4	8.6	7.7
Range	(0-29)	(2-12)	(1-23)	(3-17)
RECIPROCAL LESSON				
Average RT (Seconds)	4.7	3.0	3.4	
Percent Error	28.6	17.0	12.7	
Range	(7-51)	(5-44)	(1-18)	
INTERCEPT LESSON				
Average RT (Seconds)	9.4	7.7		
Percent Error	27.4	11.8		
Range	(4-39)	(0-24)		

Table 3 indicates that, as subjects progressed through a lesson, all recognition errors decreased and response times decreased.

Total voice recognition errors during the heading lesson declined during the four training sessions from 12.4 to 7.7 percent, but jumped to 28.6 percent and 27.4 percent during the first reciprocal and first intercept lesson, respectively. The high error rates are mainly due to nonrecognitions.

Response time (RT) is a latency measure which describes time from the appearance of an aircraft to the point at which the system accepts a command and displays feedback. The time to correct errors to achieve an input acceptable to the system is included in the RT measure.

Average response times accomplished with voice entry were examined to compare voice vs. keyboard input for component skill training. The heading lesson, in particular, was examined because data from other sources were available. Voice input response times for the heading lesson, shown in Table 3, were 6.0 and 4.6 seconds, for the first and second sessions, respectively. In comparison, keyboard entries of xxx, where xxx is a heading from 010 to 360, have been shown to produce shorter response times. A group of college students averaged 4.3 and 3.3 seconds during their first two heading lessons under very similar testing conditions (Schneider, 1988). The longer response times for voice input are due to the length of the spoken phrases in comparison to the relatively short keyboard input with average typing efficiency.

INTERSUBJECT VARIABILITY

The ranges in Table 3 indicate extreme variability between subjects. Examination of data by sex indicate that sex of subject was not a factor affecting recognition accuracy of the voice system. Overall percent error scores were 15.9 and 13.0 for the two male subjects, and 16.1, 11.1, and 18.6 for the three female subjects. No other subject variables were examined.

Table 4, showing performance and recognition accuracy scores for each subject, reveals extreme variability within subjects. Variability was defined as how well subjects maintained a consistent voice pattern from voice training to application and throughout application testing. Observations of the experimenter, reported after Table 4, suggest that difficulty of the lesson, frustration, and other emotional states may have contributed to lack of consistency. The experimenter also noted that time of day when subjects were tested may have been a factor contributing to inconsistent voice patterns for some subjects.

Table 4

Performance and Recognition Accuracy (Subject x Lesson)

	HEADING				RECIPROCAL			INTERCEPT	
SESSIONS	1	2	3	4	1	2	3	1	2
SUBJECT 1 (Female)									
Angle Error	9.0	5.5	5.4		3.9	4.5	3.4	21.4	14.7
%Correct Turns					95	98	96	88	93
%Nonrecog	23	2	1		8	41	8	2	0
%Misrecog	6	0	0		9	3	7	2	17
SUBJECT 2 (Male)									
Angle Error	4.4	3.8	4.5	4.7	5.3	4.6	3.8	12.3	13.6
%Correct Turns					91	92	94	67	100
%Nonrecog	4	2	16	4	34	4	16	27	0
%Misrecog	7	8	7	4	13	1	1	5	0
SUBJECT 3 (Male)									
Angle Error	4.8	5.2	4.2		3.5	2.7		12.2	11.3
%Correct Turns					94	94		91	100
%Nonrecog	0	2	6		21	5		32	12
%Misrecog	0	3	3		0	5		0	2
SUBJECT 4 (Female)									
Angle Error	7.7	6.5	5.9	4.4	5.6	5.2	5.3	22.1	16.3
%Correct Turns					90	91	93	75	83
%Nonrecog	9	0	1	3	6	12	1	39	24
%Misrecog	0	3	1	0	1	0	0	0	0
SUBJECT 5 (Female)									
Angle Error	7.1	7.3	6.1	5.7	9.3	8.0	4.4	17.2	14.2
%Correct Turns					84	74	79	79	81
%Nonrecog	5	8	6	6	43	6	10	18	0
%Misrecog	8	4	2	11	8	8	8	12	4

Subject 1 (Female)

Table 4 shows that recognition errors were inconsistent throughout AIC training. The unusual amount of errors, mostly rejections (41%), during the second reciprocal lesson is unexplainable. Conditions were quiet, the subject was tested in the late morning, and practice was not mentally difficult as shown by 98% accuracy on turn directions and angle error less than 5 degrees. During several of the later sessions, the subject complained about being hot and not being able to concentrate. She also said she liked the intercept lesson least of all because she could not grasp the concept. However, none of these factors seemed to be related to recognition error rates. To the experimenter, her voice sounded consistent throughout application testing. To test out the theory that boredom may affect consistency of speech patterns, this subject was given three additional heading lessons after her performance had reached criterion. The recognition error rates for these lessons were 3, 6 and 3 percent, not noticably different from the first three lessons. Time of day during which testing took place appeared to affect nonrecognition. Morning sessions averaged 18.5% rejections and afternoon sessions averaged 3% rejections. There were too few data points to perform a statistical test.

Subject 2 (Male)

Table 4 shows better recognition in the later part of training for this subject. As the experimenter heard it, this subject was not able to maintain a consistent manner of speaking. During the fourth heading lesson, the experimenter asked the subject to speak slower and consequently, the best recognition accuracy was obtained. After an unacceptable error rate during the first reciprocal lesson, the subject's voice was retrained and a new voice template was created; but this did not improve recognition. The subject showed enthusiasm and spoke loudly. Adjustment of input level may have improved nonrecognition, but Interstate does not allow manual adjustment of input level to accommodate variations in amplitudes of speakers' voices.

Subject 3 (Male)

Misrecognition errors during all lessons were less than 5%. The subject was the fastest learner of the group as can be seen by the number of sessions to reach criteria. Difficulty of learning seemed to be directly related to nonrecognition error rates. The first reciprocal lesson and the intercept lessons produced overt behaviors noted by the experimenter. The subject appeared to have a strong desire to be right and acted frustrated when he was wrong. The subject's overall manner of speaking was calm and monotone.

Background noise conditions appeared to affect system rejection of input. Its effects may have become obvious because of the fairly consistent manner of speaking in this subject, which provided a good baseline. Session 3, the third heading lesson, and session 1 of the intercept lessons, had television and voices in the background. Higher recognition error rates were obtained in these sessions, relative to other sessions which trained the same lesson. (See Table 4).

Subject 4 (Female)

Misrecognition errors during all lessons were 3% or less. The subject voiced frustrations when nonrecognitions occurred, but this did not appear to affect her voice pattern. During the two intercept lessons, learning was difficult as shown by performance accuracy scores. Average turn direction accuracy was 75% and 83%, and angle error was 22 and 16 degrees. This poor performance may have been related to the high rejection rates in the intercept lessons.

Subject 5 (Female)

This subject had high misrecognition rates and showed the poorest learning. The subject voiced frustration and disappointment with her poor performance during application testing, which to the ears of the experimenter sounded as if it affected her ability to maintain consistent voice patterns. The subject had an abnormally high number of nonrecognitions as well.

Time of day during which testing took place may have influenced nonrecognition scores for this subject. Morning sessions averaged 14.6% recognition errors and afternoon sessions averaged 4.6%, as shown in Table 4.

VOICE OUTPUT RESULTS

The Interstate/Street Electronics speech synthesizer was judged by the experimenters to be of limited usefulness for this application and so the system was not used. The speech sounded mechanical and there was a major problem with multiple word phrases running together, making the speech difficult to understand. The Votan digitization/playback unit was used to demonstrate pilot output in several field demonstrations at Navy AIC and Air Traffic Control sites. In these demonstrations, the output was judged acceptable by instructors, students, and controllers.

ADVANCED AIR INTERCEPT CONTROL TRAINING

In several field demonstrations, the testbed was used to demonstrate advanced AIC instruction using the voice recognition systems (Interstate and Votan) with one computer. In a two-plane scenario, the experimenters gave commands to a bogey and fighter via both voice input systems which activated the targets and initiated voice output (Votan). The exercise demonstrated how two trainees could train in a combat-like scenario in which two aircraft could perform evasive actions unknown to the controller beforehand. The "Enable missile" and "Fire missile" commands were included in the scenario. This form of instruction leads to high motivation because of its competitive nature.

Although voice recognition data was not collected in these demonstrations, recognition accuracy always appeared to be very good. This was, in part, due to the level of sophistication of the experimenters in using the voice systems.

SECTION V

DISCUSSION

The product of this research effort is a NAVTRASYSCEN testbed to study microcomputer system integration for application to AIC training. The testbed includes two voice recognition systems, an Interstate Vocalink 4000 standalone unit, and a Votan VPC 2000 voice card with voice output functions. The testbed demonstrates a student controller station and an instructor/pseudopilot station. The optional voice equipment at the student controller station may in some instructional situations be a viable substitute for the instructor/pseudopilot station.

MICROCOMPUTER INTEGRATION

The microcomputer air control training testbed was configured using primarily off-the-shelf equipment. All of the equipment was purchased except for the controller keyboard which was manufactured in house. It is possible to purchase a generic multifunction box to substitute for the custom-made one which would satisfy the requirements. Although the controller box was programmed to emulate NTDS air intercept control functions, with added software the box could be used to emulate several different consoles used in air control training.

Configuring the microcomputer system was extremely easy. All equipment was connected by cables. Only the Votan voice card required installation inside the computer.

The current configuration is based on an IBM PC/XT. Since the Zenith 248 is quickly becoming the standard microcomputer for Navy training, future research should consider changing to a Zenith-based system. Some of the application software was written for the TECMAR graphics board. If a generic graphics software package is written or purchased off the shelf, the application program will run on a standard Zenith or any IBM-compatible microcomputer system.

Application software, produced in a previous research effort, produced AIC training scenarios comparable to what is accomplished on the operational equipment and courseware designed to train essential visual spatial skills. No additional software development was required to interface the Interstate software with the application software. The only interface software required for the Interstate was a keyboard driver program provided by Interstate.

The hardware configuration is versatile. It can be used to test other systems requiring voice input/output at either a student or instructor station. All that is needed is compatible application software.

TRAINING ALTERNATIVES

The current configuration of the microcomputer and voice equipment provides an opportunity to research several options for training which are listed below:

1. Skill training on individual spatial components of the AIC task can be accomplished using the student station with either voice or key input. "Real-world" communications can be practiced verbally. Voice input allows the additional practice of communication skills, but will increase training time (phrases are longer) and might detract from the massive drill and practice approach required to obtain automatic skills. It can also be expected that, with voice input, recognition errors will increase as each new component skill training task is introduced.

2. Simulated whole task training can be accomplished using the student and pseudopilot stations. This configuration closely represents how the AIC schools currently conduct training. It provides realistic training in that a human with intelligence and perceptual biases hears student commands and translates them into actions. Further research is necessary to discover error rates for pseudopilots (misunderstanding of a command) and the frequency with which pseudopilots correct erroneous student commands. These data can provide a baseline with which to choose acceptable error rates for voice recognition systems, and guidance on the amount and kind of artificial intelligence required of the systems. Training systems requiring pseudopilots are undesirable because they are highly labor intensive. Often, students in the course are required to spend some of their course time in the role of pseudopilot. Research is needed to determine if there is a training value in having students in this role. In the configuration described in this report, the pseudopilot station includes a monitor slaved to the controller station. The addition of the monitor provides a passive learning situation for the trainee which may have training value.

3. Simulated whole task training using the voice input/output system at the student station is possible. The addition of the voice I/O system eliminates the need for the pseudopilot station. Testing of subjects with the Interstate voice system, however, revealed that precautions are necessary to obtain good reliability of the system. Background noise, especially talking, and the speaker's proficiency on the air control task will affect voice recognition accuracy. In addition, AIC scenarios involve aircraft flying at high speeds which requires fast and accurate responding on the part of the controllers. To get acceptable training, high accuracy rates and a minimal number of rejections must be obtained from the voice system.

The limited capabilities of the Interstate unit did not allow testing of synthesized speech output, but voice playback using Votan was successful. The small vocabulary, chosen for initial exploration of the system's potential, did not test the limits of the system. Further research is necessary to accomplish a complete demonstration of voice I/O for simulated whole task training.

4. Advanced instruction using two voice recognition systems installed at the student station and instructor station is another training alternative. Both the fighter and bogey can be activated by voice input to two recognizers. In a combat-like scenario, the moves of the bogey are unknown to the student controller and the bogey can perform evasive actions. This gives the trainee realistic training and is likely to be highly motivating because of its competitive nature.

VOICE INPUT/OUTPUT TECHNOLOGY APPLICATION

The addition of voice input/output (I/O) technology to the air control training testbed provides a research environment for developing general guidelines on the integration of voice recognition and speech generation technology with air control training devices. Several important issues can be addressed in future research: speed and accuracy requirements for various applications, appropriate forms of error correction, enrollment methods, and user training prior to using the training device. The benefits derived from the technology are highly dependent upon the specific training task, the environment, and human performance.

The performance of the Interstate Vocalink 4000 recognition system and the Votan VPC 2000 voice card in the testbed indicate several system characteristics are important to successful implementation of voice I/O. Of particular importance is the user interface during enrollment and system training, sensitivity of the voice system to background noise, and speaking habits of the user.

The greatest concern is that the voice system have a suitable user interface. If the voice enrollment and training procedures are not automated, the user is required to become skilled at the decision-making involved in what inputs should be accepted and rejected. The user learns by trial and error how vocabulary words should be joined during input to obtain the best "continuous" voice recognition. In the context of Navy training, users would be trainees. These trainees could not train the system by themselves, but would have to be supervised during voice training by an instructor who was extremely skilled in using the voice system.

Even when the system does have automated procedures for voice training, as does Interstate's Vocalink 4000, the reliability of voice recognition is still an issue. The results of the tests of voice recognition using the Interstate indicate that there are a number of factors affecting reliability. Misrecognition errors averaged under five percent in all AIC lessons; however, nonrecognitions were relatively high, ranging from 4 to 17 percent. Stress and other factors appeared to affect the speaker's ability to remain consistent. These findings are in agreement with studies that have reviewed the literature on voice recognition (Cooper, 1987). Sensitivity to speaking habits of the user is a distinct advantage, however, when part of the training task is to train consistently audible communication.

The system's sensitivity to background noise and the acceptance threshold also affected system recognition performance. With the Interstate system, background noise is easily recalibrated using a reset button on the front of the machine. The acceptance threshold, which determines how close an input must match a template in order for it to be recognized, is set by Interstate and cannot be manually controlled.

In a learning situation, the results indicate that recognition performance gets better over time as the learner becomes more proficient at the training task, but recognition performance is likely to drop again when a new task is introduced. The reasons for the drop have been related to increased mental loading, which is said to be associated with decreased recognition accuracy (Christ, 1984). A new training task may increase attention to the learning task and decrease attention to the simultaneous tasks of using correct vocabulary with the required pauses, etc. It may be necessary to retrain critical vocabulary when a new training task is introduced.

The speech output equipment used in this research allowed a comparison of speech synthesis versus recorded voice playback. The quality of voice playback was good. On the other hand, speech synthesis was difficult to program to sound acceptable to the human ear. Voice playback requires more memory than synthesis. If the size of the output vocabulary required to satisfy training requirements does not tax system storage, voice playback is preferable to speech synthesis. It may also be possible to use a combination of both playback and synthesis to reduce storage needs and get good sounding output as well.

Voice technology is continually improving. Several systems other than the ones tested in this research are on the market. Future research will undoubtedly show that voice recognition accuracy and speech synthesis can be improved over and above what was obtained in this research. Technological improvements and implementation using good human factors techniques will decrease the risks associated with implementing the technology with training devices.

SECTION VI

RECOMMENDATIONS

FUTURE RESEARCH USING THE TESTBED

The following are recommended directions for research using the microcomputer system testbed for air control training research:

1. Deliver the microcomputer configuration developed in this research effort to an air intercept control training or operational facility. Conduct evaluations of its potential uses for training.

2. Conduct further research on the cost-benefits of the pseudopilot station vs. the voice input/output system for air control training.

3. Conduct human factors research to improve the efficiency and reliability of automated voice recognition and speech generation technology applied to air control training.

4. Convert the application software to allow air control training on the Zenith 248 and several standard IBM compatible microcomputer systems.

5. Purchase a multifunction box to replace the controller box and develop software that would allow the box to be programmed to emulate several different operational equipment consoles.

6. Research the possibility of writing automated scenario generation software that would produce scenarios and component skill lessons for several different types of air control training.

7. Develop other application software that would allow the present hardware configuration to be used to test voice input and output in other applications.

USE OF VOICE INPUT/OUTPUT TECHNOLOGY WITH TRAINING DEVICES

The following are recommendations for pre-enrollment training, enrollment training, and use of the recognition system during technical training, as they pertain to air control training devices:

Pre-Enrollment Training

1. Train users to regularly check and adjust microphone position.

2. Train users to regularly adjust input level control to accommodate various amplitudes of speakers' voices.

3. Present a written copy of the vocabulary.

4. Present verbal instructions to point out necessary syntax and vocabulary constraints, i.e., forced pauses, limits on length of input, etc.

5. Expose the trainee to a sample scenario showing how the vocabulary would be used and allow the trainee to speak in the scenario.

6. Give verbal feedback on speaking style to improve the trainee's ability to remain consistent throughout use of the system.

Enrollment and System Training

1. Train utterances in noisy and quiet conditions, if both noisy and quiet conditions will occur during use.

2. Use the same type of microphone that will be used with the training device.

3. Train utterances under stress conditions similar to those that will apply during use.

4. Prompt with text (visual) not words (auditory) for beginning enrollment.

5. Use actual controller transmissions in the recognition test after voice training.

6. Introduce variability into the speech samples to widen the range of acceptable responses during use.

7. Feedback on recognition accuracy should not exceed or differ from what will occur during use.

Application: Air Control Training

1. Use voice input when the training goal is to train air control vocabulary, and keyboard input when training visual-spatial skills or other skills.

2. Retrain utterances as needed to accommodate changes in speech patterns as trainees are introduced to new air control training tasks.

3. Do not have air control training sessions last more than one hour.

4. Retrain one or two critical vocabulary items routinely each time the system is to be used to accommodate short-term variations in speech characteristics or environmental noise.

5. Design noticable feedback and easy correction procedures to reduce problems when timing is critical to the technical task and trainees are beginning users.

6. Regularly check and adjust microphone position.

7. If an utterance is misrecognized several times throughout air control training, retrain the subject on that utterance.

8. Monitor background noise and design unobtrusive adjustment procedures if the background noise changes significantly during a training session.

9. Design the training environment to minimize the interference produced when trainees are talking in close proximity to each other. Use noise-cancelling microphones and soundproof materials between stations when possible.

REFERENCES

Christ, K. A. (1984). Literature review of voice recognition and generation technology for Army helicopter applications. (Technical Note 11-84). Aberdeen Proving Ground, MD: U.S. Army Engineering Laboratory. (NTIS No. AD 146 878)

Cooper, M. (1987). Human factors aspects of voice input/output technology. Speech Technology, March/April, (pp. 82-85). New York, NY: Speech Technology, Inc.

Halley, R. (1981). Training effectiveness report for air intercept controller prototype training system. (Technical report 78-C-0182-12). Orlando, FL: Naval Training Equipment Center.

McCauley, M. E. (1984). Human factors in voice technology. In Muckler, F. A. (Ed.). Human Factors Review 1984 (pp. 131-166). Santa Monica, CA: Human Factors Society.

Poock, G. K. (1986). Speech recognition: Research, applications, and international efforts. Proceedings of the Human Factors Society 30th Annual Meeting, Volume 2, 1278-1283.

Schneider, W. (1982). Automatic/control processing concepts and their implications for the training of skills. (Technical Report 8101). Champaign, IL: University of Illinois, Human Attention Research Laboratory.

Schneider, W. (1988). Developing automatic component skills for air intercept control. Interim Technical Report, Pittsburgh, PA: Learning Research and Development Center, unpublished.

Skriver, C. P. (1987). Airborne message entry by voice recognition. Proceedings of the Human Factors Society 31st Annual Meeting, Volume 1, 424-427.

APPENDIX A

INTERSTATE VOCABULARY LIST AND SYNTAX

The vocabulary used in the Air Intercept Control training application is shown below in the vocabulary list. The syntax file, which identifies word orders acceptable to the system, is also presented. A slash identifies an optional word, KEY defines a category of first words, and FIN defines a category of last words. A graphic representation of the syntax structure is presented in Figure A-1.

VOCABULARY LIST

enable missile
fire missile
viper
snake
0
1
2
3
4
5
6
7
8
9
for the bogey
as bogey
left
right
heading

SYNTAX FILE

```
.KEY .DIR/ .DIGIT1 .DIGIT ZERO .FIN/
ENABLE-MISSILE
FIRE-MISSILE

.FIN=
FOR-THE-BOGEY
AS-BOGEY

.KEY=
VIPER
SNAKE
HEADING

.DIR=
LEFT
RIGHT

.DIGIT1=
         0
         1
         2
         3

.DIGIT=  0
         1
         2
         3
         4
         5
         6
         7
         8
         9
```

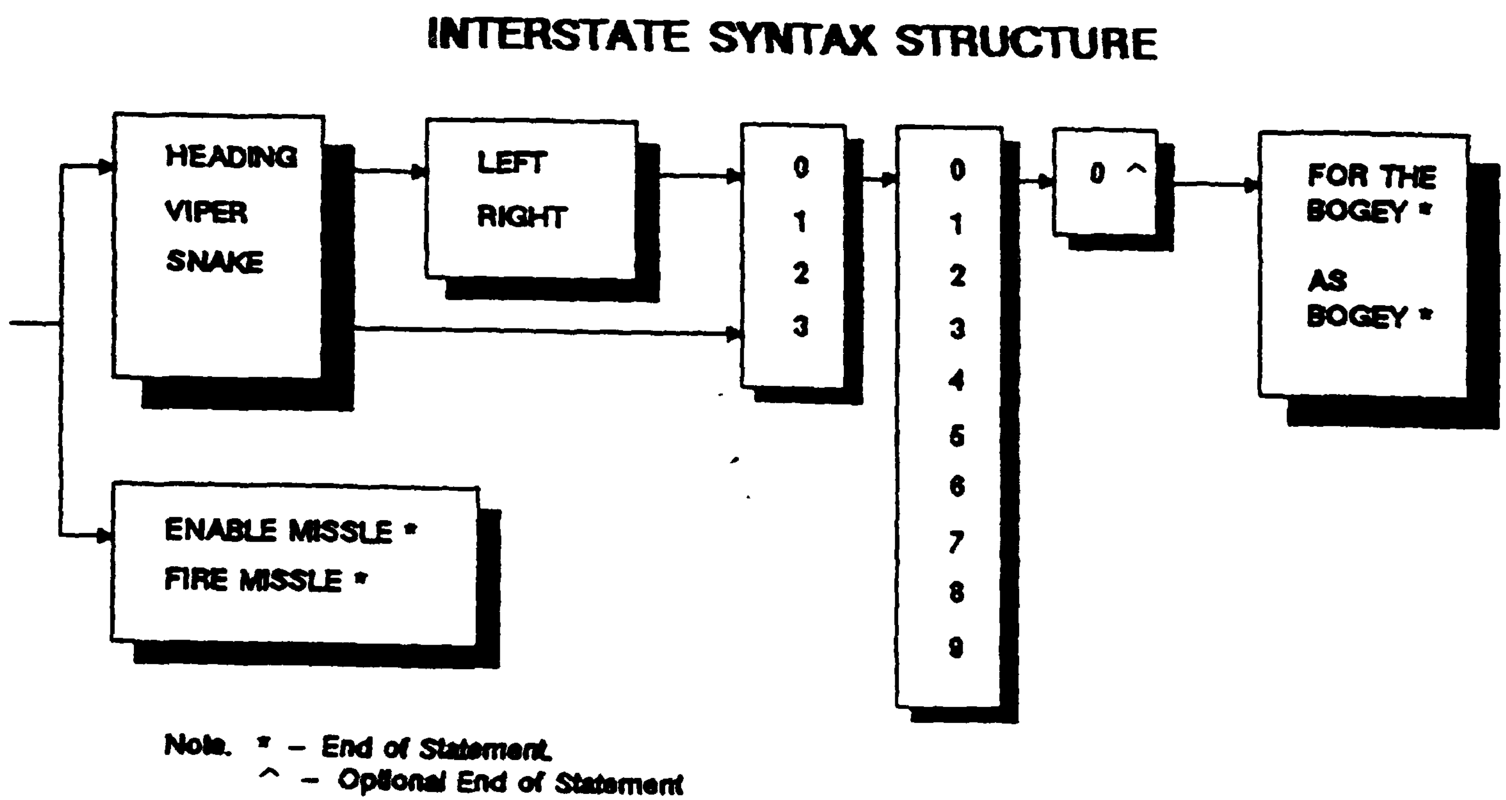


Figure A-1. Interstate syntax structure for the air intercept control vocabulary.

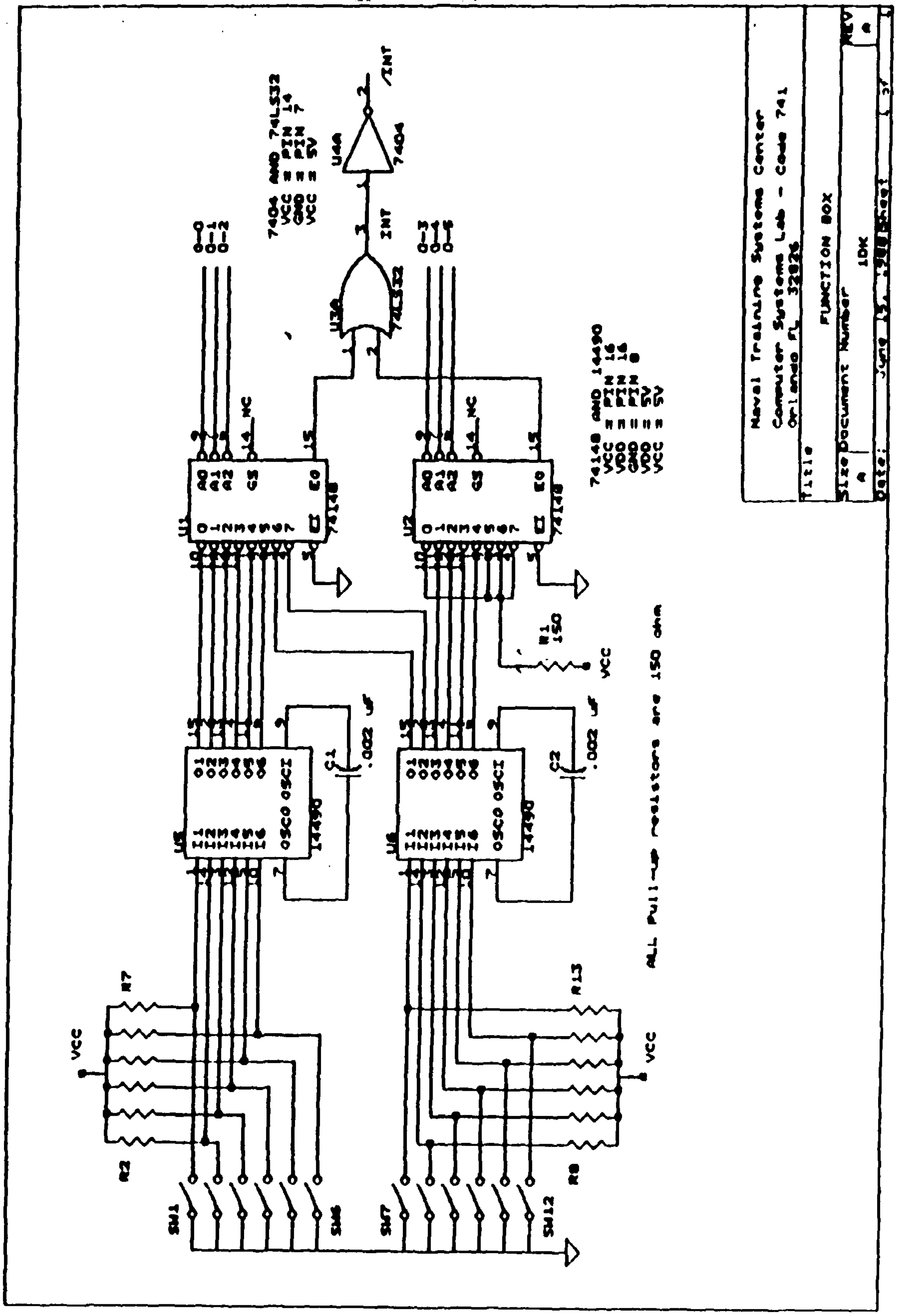


Figure B-1. Schematic drawing of the controller box circuit card.

DISTRIBUTION LIST

Commanding Officer
Naval Aerospace Medical Institute
Naval Air Station
Code OOL
Pensacola, FL 32508-5600

Commander
Naval Air Systems Command
Naval Air Systems Command Headquarters
AIR 950D - Technical Library
Washington, DC 20361-0001

Commander
Naval Air Systems Command
Naval Air Systems Command Headquarters
Code 933G
ATTN: CDR Steve Harris
Washington, DC 20361-0001

Commander
Naval Air Systems Command
Naval Air Systems Command Headquarters
Code PMA 205
ATTN: LT James Hooper
Washington, DC 20361-0001

Commander
Naval Air Systems Command
Naval Air Systems Command Headquarters
Code PMA 205
ATTN: CDR Jerry Owens
Washington, DC 20361-0001

Commander
Naval Ocean Systems Center
Code 441
San Diego, CA 92152-5000

Commanding Officer
Naval Research Laboratory
Library
Washington, DC 20375

Commander, Naval Sea Systems Command
Naval Sea Systems Command Headquarters
CEL-MP
ATTN: Dr. Judy Moracco
Washington, DC 20362-5101

Chief of Naval Education and Training
NAS, Pensacola, FL 32500-5100

Commanding Officer
Naval Training Systems Center
Code PM34
Attn: Col Rounseville
Orlando, FL 32826-3224

Commanding Officer
Naval Training Systems Center
Code 1
Orlando, FL 32826-3224

Commanding Officer
Naval Training Systems Center
Code 7
Orlando, FL 32826-3224

Commanding Officer
Naval Training Systems Center
Code 7A
Orlando, FL 32826-3224

Commanding Officer
Naval Training Systems Center
Code 71
Orlando, FL 32826-3224

Commanding Officer
Naval Training Systems Center
Code 73
Orlando, FL 32826-3224

Commanding Officer
Naval Training Systems Center
Code 74
Orlando, FL 32826-3224

Commanding Officer
Navy Personnel Research and Development Center
Code 01
ATTN: Dr. J. McMichael
San Diego, CA 92152-6800

Commanding Officer
Navy Personnel Research and Development Center
Code 51
ATTN: Dr. E. Aiken
San Diego, CA 92152-6800

Commanding Officer
Navy Personnel Research and Development Center
Code 52
ATTN: Dr. J. McLachlan
San Diego, CA 92152-6800

Department of the Navy
Office of the Chief of Naval Operations
OP-01B2
ATTN: Dr. R. Carroll
Washington, DC 20350-2000

Department of the Navy
Office of the Chief of Naval Operations
OP-987H
ATTN: Dr. Bart Kuhn
Washington, DC 20350-2000

Chief of Naval Research
Office of the Chief of Naval Research
Code 1142PT
ATTN: Dr. Susan Chipman
Arlington, VA 22217-5000

Naval Medical R&D Command
NMRDC-404
NMC NCR Complex
ATTN: CAPT Thomas Jones
Bethesda, MD 20814-5044

Chief of Naval Research
Office of the Chief of Naval Research
Code 1142 EP
ATTN: Dr. John O'Hare
Arlington, VA 22217-5000

Office of Naval Technology
Code 222
ATTN: Dr. Stan Collyer
800 North Quincy Street
Arlington, VA 22217-5000

Chief, U.S. Army Research Institute
Orlando Field Unit
ATTN: Dr. H. Ozkapton
12350 Research Parkway
Orlando, FL 32826

HQDA
DAPE-ZXO
ATTN: Dr. M. Fischel
The Pentagon
Washington, DC 20310-0300

PMTRADE
AMPCM-AVD
ATTN: Col Lunsford
Orlando, FL 32826

Chief, U.S. Army Research Institute
ATTN: Dr. Edgar Johnson
5001 Eisenhower Avenue
Alexandria, VA 22333

Chief, U.S. Army Research Institute
ATTN: Dr. Hiller
5001 Eisenhower Avenue
Alexandria, VA 22333

AFAMRL/HEF
Director, Human Engineering Division
Wright-Patterson AFB, OH 45433

AFHRL
Training Systems
ATTN: Col W. J. Lobbestael
Wright-Patterson AFB, OH 45433

AFHRL
Operations Training Division
ATTN: Col Michael Lane
Williams AFB, AZ 85240

Air Force Office of Scientific Research
Technical Library
Bolling AFB, DC 20319

AFHRL/OTT
ATTN: Dr. Dee Andrews
Williams AFB, AZ 85240-6457

Dr. John Forester
AFAMRL/HEC
Bldg 248, Area B
Wright Patterson AFB, OH 45433-6583

American Psychology Association
Psyc. Info. Document Control Unit
1200 Seventeenth Street
Washington, DC 20036

Defense Technical Information Center
FDAC
ATTN: J. E. Cundiff
Cameron Station
Alexandria, VA 22304-6145

Headquarters
U.S. Marine Corps
Scientific Advisor
Washington, DC 20380

Marine Corps Program Manager
for Training Systems
Code SST
Marine Corps Research, Development
and Acquisition Command
Quantico, VA 22134

Institute of Defense Analysis
Science and Technology Division
ATTN: Dr. Jesse Orlansky
1801 Beauregard Street
Arlington, VA 22311

National Defense Institute
Research Directorate
Ft McNair, DC 20319

OUSDR&E (E&LS)
ATTN: Dr. Earl Alluisi
The Pentagon
Washington, DC 20301-3080