REPORT NAVTRAEQUIPCEN 77-C-0162-6

12

LEVEL III

AD A092719

GROUND CONTROLLED APPROACH CONTROLLER TRAINING SYSTEM
FINAL TECHNICAL REPORT

A091930

MARY HICKLIN, GENE BARBER, JOHN BOLLENBACHER,
MICHAEL GRADY, DAVID HARRY, CATHERINE MEYN,
GAIL SLEMON
LOGICON, INC.
TACTICAL AND TRAINING SYSTEMS DIVISION
POST OFFICE BOX 80158
SAN DIEGO, CALIFORNIA 92138

April 1980

FINAL REPORT FOR PERIOD 23 SEPTEMBER 1977 – 23 APRIL 1980

DOD DISTRIBUTION STATEMENT

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

NAVAL TRAINING EQUIPMENT CENTER
ORLANDO, FLORIDA 32813

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NAVTRAEQUIPCEN 77-C-0162-6	2. GOVT ACCESSION NO. AD-A092 719	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) GROUND CONTROLLED APPROACH CONTROLLER TRAINING SYSTEM (GCA-CTS) FINAL TECHNICAL REPORT		5. TYPE OF REPORT & PERIOD COVERED Final Report 23 SEP 1977 - 23 APR 1980
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Mary Hicklin, Gene Barber, John Bollenbacher, Michael Grady, David Harry, Catherine Meyn, Gail Siemon		8. CONTRACT OR GRANT NUMBER(s) N61339-77-C-0162
9. PERFORMING ORGANIZATION NAME AND ADDRESS Logicon, Inc. P.O. Box 80158 San Diego, California 92138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 7796-1P1
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Training Equipment Center Orlando, Florida 32813		12. REPORT DATE April 1980
		13. NUMBER OF PAGES 112
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

VTAG (Voice Technical Advisory Group)
Automated Adaptive Training
Automated Speech Technology
Ground Controlled Approach Controller Training System
Isolated Phrase Recognition
Speech Recognition
Speech Synthesis
Speech Understanding
Voice Data Collection

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

This report describes the development of an experimental prototype Ground Controlled Approach Controller Training System, an automated adaptive training system designed to teach precision approach radar control skills. The system is unique in that it is the first to employ the automated speech technologies to automate the training of a primarily verbal task. The design embodies all the lessons learned in previous laboratory studies. This report outlines the rationale behind the system design and offers observations, conclusions and recommendations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406012

FOREWORD

This report provides a review of the design, integration, installation, and check-out of a precision approach radar experimental prototype training system. Computer speech understanding technology was tested within the real-world training curricula at the Navy's Air Traffic Control School for the precision radar task.

This is the first military training application of computer speech understanding and this report is intended to document the experimental nature of the effort. It will be of interest to speech researchers for its discussion of real-world problems, to systems analysts for its integration lessons-learned, and to training analysts for its implications for training system design of speech tasks.

The use of speech understanding technology is expected to contribute to the solution of manpower shortage problems, especially when speech technology is combined with other advanced training technologies such as automated performance measurement, feedback, critique and diagnosis. Applications for other tasks are under consideration.

R. BREAUX, Ph.D.
Scientific Officer

ACKNOWLEDGMENT

The project was fortunate to have the guidance, encouragement and support of its Scientific Officer, Dr. Robert Breaux. Dr. Breaux, a training psychologist, has guided the investigation of the application of the speech technologies to training from laboratory studies through this experimental prototype. His vision and interest in the project have contributed enormously to its success.

Military and Government personnel at the Naval Air Technical Training Center, Millington, Tennessee, have also contributed significantly to the project. CWO Robert Lynchard has been uniformly supportive throughout the project. His predecessor, ACCM Jim Beck, also gave generously of his time and suggestions. LDO Gary Tate and ACC Walter Hall served as subject matter experts during the initial phase of the project. ACC Gerald H. Cyr and AC1 Terry M. Martin devoted an enormous amount of effort to the project serving as subject matter experts and participating as learning supervisors to GCA-CTS users. They have contributed a great many suggestions which reflect their keen interest in the project. John Norrell of Education and Training Support Training Device has helped us on numerous occasions. Finally, computer center personnel have been extremely cooperative in making the system available to us at odd hours.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (CONT)

LIST OF TABLES

LIST OF TABLES (CONT)

SECTION I

INTRODUCTION

The experimental prototype version of the Ground Controlled Approach Controller Training System (GCA-CTS) represents the culmination of work begun in 1972 for the Naval Training Equipment Center (NAVTRAEQUIPCEN). A principal concern of this agency is the identification and capture of those quantifiable aspects of human behavior that relate to the improvement of performance through training. This focus requires a continual scan of modern technologies to spot developments which can be applied to training systems. The observation that there exists a class of job situations which have in common the use of restricted, stylized speech, and the fact that studies have shown that it has been possible to achieve savings of manpower and training time while joining uniform, high-quality training through the use of automated adaptive instruction (Goldstein, Norman, et al., 1974) led to a step-by-step investigation of the application of the emerging speech recognition technology to military training.

The earliest work identified the Precision Approach Radar (PAR) control task as an ideal test bed for research in this area because it was a primarily verbal task not previously amenable to automated training and because the vocabulary used was rigidly defined and highly stylized and hence, potentially recognizable by the isolated phrase recognition technology. This work also established the conceptual feasibility of an automated adaptive GCA controller training system (Feuge, Charles and Miller, 1973).

Subsequent work was devoted to demonstrating the technical feasibility of the concept in a series of laboratory studies which involved the development of a preliminary training system. This laboratory version demonstrated the feasibility of real-time speech understanding of the PAR vocabulary, and of performance measurement and adaptive syllabus control for this primarily verbal task. In addition, a sophisticated application-independent voice data collection technique was developed for the laboratory system and shown to be feasible (Grady and Hicklin, 1976).

The present project, begun in 1977, was intended to bring the technology out of the rarefied laboratory atmosphere and into the operational training environment as a stand-alone, fully automated adaptive training system. This goal led to the development of a totally new application-oriented system which was firmly grounded in the previous development work and which took cognizance of the research results obtained using the laboratory system (e.g., Breaux, 1976). Particular attention has been devoted to the attainment of user acceptance in order to maximize the usefulness of the training system (Hicklin and Slemon, 1978).

The goals of the Ground Controlled Approach Controller Training System project, then, have been to develop an experimental prototype training system which would demonstrate the feasibility of:

a. Employing the automated speech technologies in an operational training environment;

b. Developing a training methodology in conformance to the philosophy of instructorless training and to the limitations of the automated speech t chnologies without compromising training effectiveness;

c. Developing an instructor model which could provide automated adaptive training for a primarily verbal task;

d. Devising a performance measurement scheme which would enable the system to provide instructive feedback to the trainee, informative progress information to the learning supervisor, and input to the instructor model which would enable automated adaptive problem selection;

e. Devising techniques for providing the feedback to the trainee and learning supervisor;

f. Developing useful models of the verbal and motor behavior of the other persons with whom the precision approach radar controller interacts, namely the pilot, pattern controller and tower controller, as well as a model of PAR controller behavior.

The primary constraints which shaped the GCA-CTS involved:

a. The need to utilize the state-of-the-art speech recognition technology which was exemplified at the hardware procurement stage of the project in the Threshold 500 isolated phrase recognition device;

b. The need to achieve good speech understanding in real time over a relatively large vocabulary containing many similar phrases;

c. The need for the system to be usable by persons not previously trained in the use of speech recognition equipment;

d. The need for the system to provide training in the PAR control task equivalent to that provided in the existing training environment, in an environment with a minimum of instructor intervention;

e. The need for the system to provide realistic stimuli (e.g., radar display, servo control, communications gear, etc.) to facilitate transfer of training;

f. The need for the system to be able to train the student to proficiency in eight days. During the course of the development, the school's training time was reduced to five days, and the GCA-CTS courseware was modified to account for this reduction. After courseware development was substantially complete, the system time was reduced again to five half days per student so that two students could use the system in a week.

OVERVIEW: AN EXECUTIVE SUMMARY

The fc .' of the previous laboratory system was on the development of a tool for studying general questions about the application of speech recognition to training. The speech recognition subsystem itself was designed to be completely application independent to facilitate this type of research. Since a speaker-dependent recognition capability requires that samples of the individual talker's voice be on file before recognition is possible, a flexible, stand-alone voice data collection capability was also developed to investigate techniques for automatic collection of representative voice patterns (Breaux and Grady, 1976). Existing hardware and software were incorporated wherever possible to minimize the cost in nonessential areas. Thus, an existing aircraft/pilot model implemented on NAVTRAEQUIPCEN's PDP-9 was used, as was the graphics package and graphics monitor available on that system. Furthermore, no thorough GCA task analysis was performed. The performance measurement system was primarily intended to demonstrate the classes of performance which could be monitored and evaluated.

The purpose of the present project was to develop a prototype which (quoting from the contract specification) "will be used to fine tune the subsystems of the GCA-CTS and validate the successful laboratory evaluation." The focus of the prototype has therefore been upon the development of a training system. To devise such a system, it was necessary to compromise some of the application independence of the laboratory subsystems and to develop entirely new subsystems as well, integrating them into a system tailored precisely to meet the needs of the specific application.

One of the first project tasks was the identification of behavioral objectives for the GCA controller task. These objectives were developed on the basis of the task analysis work reflected in Training Characteristics of the Automated Adaptive Ground Controlled Approach Radar Controller Training System (Breaux, 1976), and on the basis of discussions with subject matter experts at the Naval Air Technical Training Center (NATTC), Millington, Tennessee.

Early in the project, the ideal of an "instructorless" training system, which would take over all of the routine duties of the instructor emerged. Such a system would free him or her to perform the learning supervisor tasks so often neglected for want of time. This system concept implied that the GCA-CTS must not only automatically and adaptively provide practice problems and provide feedback as the laboratory system had done, but also must provide a course of instruction into which voice data collection was smoothly incorporated.

The results of the task analysis, together with the constraints imposed by the instructorless training system concept and by the limitations inherent in the state-of-the-art technology which was being employed posed challenging problems in the areas of software, courseware and hardware system design. Software design requirements included the need for a computer-managed instructional system which would provide the courseware designer with a flexible, useful tool. In addition, software enhancements to state-of-the-art speech rec[illegible]on capability were essential. Courseware design problems included the [illegible]d to teach the task in an instructorless environment while minimizing the impact of the limitations of the technology on training. Hardware design problems involved devising ways of monitoring diverse aspects of student behavior such as speech level, servo manipulation and transmitter keying. They also included devising techniques to provide adequate feedback to the trainee about his verbal performance.

The emerging system concept was gradually refined and from it, four phases of instruction were defined: interactive teaching, commented practice, graded practice and performance test. A hierarchical table structure was devised to control the course of instruction such that the order and content of these phases could be varied during courseware development without correlative software changes.

An "errorless learning" philosophy was determined to be consonant with the apparently formidable constraints placed upon the courseware by the instructorless training concept and by the requirement to collect from four to ten samples of each of the more than 100 phrases used by the GCA controller. Briefly, the concept underlying the development of the course syllabus was that the material could be partitioned into the right-sized chunks such that in learning it, the trainee would never learn to make errors. Further, while the trainee was learning to use a particular radio transmission, speech samples could be collected. Thus, the potentially grueling voice data collection procedure would be spread out over the entire training period. Furthermore, the speech samples, elicited in a simulated control situation, were more likely to be representative of the trainee's normal voicing than are speech samples elicited by the directive to repeat a phrase ten times in succession. Perhaps most importantly, the concept allows the trainee to develop confidence in the speech recognition system while it is performing recognition over a subset of the very large vocabulary when the potential for good speech recognition for a naïve user is greatly enhanced.

Hardware design solutions required the design of special purpose equipment for motor behavior monitoring and for digital speech recording and playback.

As a result of the implementation of these design concepts, the GCA-CTS is a stand-alone, experimental prototype training system which provides automated, individualized instruction in the techniques applicable to providing ground-controlled approaches. In addition, it provides a realistic environment in which the radar control skills can be practiced under the supervision of an automated instructor which provides objective performance measurement and feedback in the form of performance summaries and annotated replays. Although the order of topic presentation is rigidly defined in the basic syllabus, problem difficulty is adapted, amount of practice is varied, and remedial exercises are selected to automatically adapt the basic course to the needs of the individual trainee. One of the major benefits of the system is that it relieves the trainee of the need to devote part of his or her time to serving as a pseudo pilot for other trainees — a requirement when using the existing training device. It also provides enrichment topics for those students who complete the basic course quickly. Finally, the system provides the learning supervisor with informative feedback about the individual trainee's performance.

Results presented in this report suggest that the system concept is a viable one. Based upon a small sample of trainees, it seems apparent that the use of the speech recognition system can be taught to naive users during a short period of time such that recognition accuracy sufficient to support training can be achieved.

BACKGROUND: HOW PROJECT REQUIREMENTS WERE MET

Table 1 lists the GCA-CTS project deliverables. The following discussion briefly describes how these project requirements were met.

TABLE 1. CGA-CTS DELIVERABLES

ITEM	DESCRIPTION	DELIVERED
0002	Work Plan Report	10-23-77
0003	Quarterly Progress Reports	every 90 days
0004	Training/Functional Design Report	2-23-78
0005	System Configuration/Facilities Report	9-23-78
0006	Demonstration Test Plan	2-23-79
0007	Trainer Demonstration Report	9-23-79
0008	Training System Package	8-23-79
0009	Training Effectiveness Test Plan	5-23-79
0010	Training Effectiveness Testing	
0011	Interim Support and On-Call Service	
0012	Preliminary Technical Reports	2-23-80
0013	Final Documentation	2-23-80
0014	Summary Oral Briefing	4-11-80
0015	Functional Specification	2-23-80
0016	Final Technical Report	4-23-80

WORK PLAN REPORT. A project work plan was prepared and submitted to the Government for review as line item 0002 of the contract. The report was revised and resubmitted in December, 1977.

The work plan report was prepared as a useful working notebook, utilizing a loose-leaf binder format with sections for the contract, project deliverables, task statements, correspondence, etc.

The report was discussed and reviewed at an orientation meeting held in November, 1977, at the Naval Air Technical Training Center, Memphis, Tennessee. he meeting was attended by personnel from the Naval Air Systems Command (NAVAIRSYSCOM), NATTC, Naval Education and Training Program Development Center (NAVEDPRODEVCEN), Naval Training Equipment Center (NAVTRAEQUIPCEN) and Logicon.

QUARTERLY PROGRESS REPORTS. Quarterly reports were prepared throughout the course of the project which described work completed during the quarter, work in progress, problem areas, and project resources expended and remaining as well as actual and projected expenditures.

TRAINING/FUNCTIONAL DESIGN REPORT. This document served as the baseline for the GCA-CTS design insofar as it detailed what the experimental prototype system had to be capable of doing. The report consisted of several sections and appendices.

SYSTEM CONFIGURATION/FACILITIES REPORT. This report documented the hardware and software environment, the design of the special purpose hardware including drawings, and the software design of the GCA-CTS including data definitions and file structures. It also presented the results of a survey of the facilities at NATTC and indicated the required modifications to the existing facilities.

DEMONSTRATION TEST PLAN AND REPORT. The Demonstration Test Plan (Logicon, 1979a) provided a set of tests to demonstrate the GCA-CTS capabilities. The results of the testing were appended to the original document and the resulting document was published as the Demonstration Test Report (Logicon, 1979b).

TRAINING SYSTEM PACKAGE. Implementation of the system involved integrating the Government-furnished equipment (GFE) and vendor-supplied hardware, installing the special purpose hardware, courseware development, software development, and document preparation.

TRAINING EFFECTIVENESS. A Training Effectiveness Test Plan was devised and the results are described in the present document.

INTERIM SUPPORT AND ON-CALL SERVICE. The system has been supported by on-site personnel. Requested enhancements have been added.

FINAL DOCUMENTATION. Final revisions of the Student Guide (Hicklin, et al., 1980b), Instructor Guide (Hicklin et al., 1980a), and System Documentation (Barber et al., 1980) have been prepared. The present document completes the set of required final documentation.

ORAL BRIEFING. A summary briefing was presented at the Naval Training Equipment Center on April 11, 1980.

FUNCTIONAL SPECIFICATION. A functional specification for the retrofit of the existing training device, the 15G19, with GCA-CTS-like capabilities has been prepared.

FINAL TECHNICAL REPORT. The present volume is the final GCA-CTS deliverable. It provides an historical review of the project along with conclusions and recommendations. The subject matter outline conforms to that specified in NAVTRAEQUIPCEN specification DI-E-2119/MOD.

ASSUMPTIONS

This development effort was based upon certain instructional and technological assumptions which are discussed below.

INSTRUCTIONAL ASSUMPTIONS. The underlying assumption which governed the development of the GCA-CTS was that it was possible to provide automated adaptive training for this primarily verbal task. The most significant implication of this assumption from the instructional perspective is that it must be possible to devise a user acceptable system. It was reasoned that if the system proves difficult to use, irritating to listen to, or fails consistently to recognize the spoken advisories, the trainee's frustration would probably impede learning. Implied in the underlying assumption, then, is the assumption that the user acceptance risks could be adequately overcome, and a concern for user acceptance pervaded the system design. The instructional assumptions listed below are specifications of the underlying assumption. They are:

a. That a training course could be developed which complements the technology being employed;

b. That the system could automatically teach the student how to use itself;

c. That an automated instructor model could be developed;

d. That a performance measurement scheme could be devised in which pilot performance does not impact the student's grade;

e. That effective feedback could be devised;

f. That stylization constraints would not have a negative impact on transfer of training or user acceptance;

g. That a Student Guide should serve as the primary source of instructional material;

h. That realism in the simulations was required, but such things as a look-alike console were not.

A Training Course Which Complements the Automated Speech Technology. There are many possible approaches to teaching GCA controller skills. The present (non-automated) course requires that the trainee master the entire vocabulary and try to put it to use the first time he plays the role of final controller. At first, the instructor stands by and prompts him, and the pseudo pilot (a trainee also) can easily understand the radio transmissions despite stylization problems and even word substitution errors. During the five days of training, errors decrease dramatically and the trainee emerges as a proficient controller.

This training philosophy is not suitable for an automated training system, however. The automated system has certain strengths and certain limitations when compared to the human pseudo pilot and human instructor. The automated system cannot match the verbalization error tolerance of the human listener for example, but it can be much more attentive than an instructor who is responsible for several students. It has the flexibility to simplify the environment and even stop the approach if necessary to illustrate its points. It can also contrive practice approaches which require use of only that material learned to date. For example, until the glidepath transmissions are learned, the glidepath display is not shown on the indicator.

To take advantage of the tremendous power of the automated training system and to minimize the impact of its shortcomings, a training course was designed which was assumed to maximize the training effectiveness of the GCA-CTS. This training is in accordance with the principles of errorless learning and the ideal is for the trainee never to learn to make mistakes. Furthermore, this training strategy actually takes advantage of what is often considered to be a drawback in state-of-the-art speaker dependent speech recognition, namely the requirement to configure the system for the individual's speech patterns. Briefly, the training strategy involves dividing the GCA controller task into its component parts so that one topic can be presented at a time. The simulated environment is manipulated to provide illustrative examples while the instructor simulation provides prompts. The trainee's speech patterns are collected at this time and used to create the reference patterns for speech recognition. Thus, as the trainee learns the procedures and phraseology, the system is learning to recognize his voice. After this phase, the system presents tasks for the trainee to perform without prompts. These are scored. Their purpose is to allow the student to practice the new material and integrate it with the old.

This training strategy was expected to impact user acceptance in several ways. First, use of the system should prove rewarding. Initial speech recognition problems are minimized by the small vocabulary and thorough training so the goals set forth are readily attainable. This initial success bolsters the trainee's confidence in the system, thus providing an important ingredient for future recognition success.

Secondly, the formidable task of configuring the system to recognize the large GCA-CTS vocabulary has been integrated with task training in a way that is transparent to the user. User acceptance is not hindered by an en masse onslaught of phrase repetition requests. In addition, the strategy employed in speech pattern collection ensures representative reference patterns and therefore further promotes speech recognition accuracy.

Finally, the automated adaptive system devotes all of its resources to teaching the individual user. It ensures that each topic is mastered in a systematic way at the trainee's own pace. Any needed remediation is provided automatically. Because of this highly individualized instruction, the trainee is expected to attain proficiency in all of the GCA control tasks. The proficiency level attainable through thorough, systematic, task-oriented learning can be its own reward and enhances user acceptance.

Training GCA-CTS Users. Proper use of the training system is the first topic addressed in the training program. The rules of microphone placement and speech level production are critical to good speech recognition, but are easy to learn and employ. There is also a less obvious component of proper system use which has been labeled "learning to talk to the box." Introspection, if it may be admitted, suggests that the components of this art include confidence, naturalness of speech and consistency. The skill is easily acquired, yet time for acclimatization is very important to good speech recognition. It was assumed that the system could teach these skills through providing explanations and practice situations. During the practice period, the trainee is encouraged to experiment with the system to learn that it really can recognize what he or she says, and to discover the limitations inherent in automatic speech recognition. With this background, the trainee is ready to use the system effectively.

Automated Instructor. It was assumed that an instructor model could be devised which would provide instruction and automatically tailor the course to the individual trainee's needs based upon performance data.

Performance Measurement. It was assumed that it would be possible to quantify and model the performance assessment measures used by qualified instructors in a way which would allow error detection and reporting and adaptive problem selection. One aspect of concern, based upon the laboratory studies, was to prevent pilot behavior from affecting the assessment of trainee performance because the system is, after all, intended to teach controllers, not pilots. This focus was maintained throughout the process of extracting the behavioral objectives and the specifying of the performance measurement variables. An exception to this rule was made in assessing the quality of the initial turn to the final approach heading; but in all other cases, the performance assessment is independent of the quality of the pilot's simulated motor response to control instructions. Furthermore, a flexible scoring algorithm was developed which makes it possible for any skill category (such as turns to final) to be omitted from consideration in scoring simply by specifying this election in the problem specification file.

Effective Feedback. Despite the precautions taken to ensure errorless learning, consistently error-free performance in the complex GCA controller environment will probably remain a lofty ideal which can only be imperfectly realized. Therefore, what was assumed to be effective feedback was designed to enable the user to understand and learn from his mistakes. In this unique training environment, these mistakes include stylization errors which cause recognition failures. The stylization error feedback used in the laboratory system proved to be less than ideal. The replay capability added to the laboratory GCA-CTS made use of the existing speech synthesizer to repeat the trainee's advisories and give rule explanations when an error was encountered. It was noted that "message not understood" reports proved especially frustrating to the students. Quoting from a memo Dr. Breaux presented to a project review meeting on January 10, 1978: "A significant aspect to the capability of GCA-CTS training is user acceptance. This can be readily destroyed by 'message not understood.' Now 100 percent recognition accuracy may not be required, totally, if there is proper feedback from the system." Conversations with instructors at NATTC in November, 1977 who had used the laboratory system revealed that their "largest complaint with the laboratory system was that the user couldn't argue with the system." Many times a user was convinced that he had uttered the correct advisory but he had no way to argue with the computer or to understand why the recognition failure occurred. For example, he would remember that he had said "slightly above glidepath" but would not realize he had paused in mid-phrase, making the transmission unintelligible to speech recognition.

To add to replay's usefulness as a feedback technique and to reduce the frustrations associated with recognition failures, a speech input digitizer was designed to record the trainee's advisories. The replay then consists of an actual recording of the trainee's speech, synchronized with the visual display. From this, the student should be able to understand the cause of any recognition failures which occur. In addition to enhancing the student's acceptance of training system decisions, the replay provides the instructor with an excellent tool. After the run the student and instructor can review the run together. This provides an excellent forum for discussion of such things as subtle points of style.

Other forms of feedback were also devised. In the optional commented practice phase 2 mode, the system stops and explains if an error is made on the new material. Laboratory experience showed that this is disruptive to a train of thought, so the system restarts each problem after an error is detected and explained.

Stylization Constraints. Since the GCA-CTS employs an isolated phrase recognition capability, the trainee must conform to three important stylization constraints to enable recognition to occur. These are: 1) proper phraseology must be used; 2) a slight pause must be inserted after each phrase; and 3) pauses must not be inserted in the middle of phrases. It was hypothesized that none of these constraints would adversely affect training. In fact, they may actually augment training. The trainee quickly learns that the system will not recognize non-standard phraseology, so he never gets in the habit of using it. The slight pauses required between phrases actually encourage the novice to slow down and think about what is to be spoken, then utter it confidently and carefully. The instructors supported this feature because they have reported that often the inexperienced controller gives control information more rapidly than the pilot is able to respond to it, and is not sufficiently careful about ensuring that the information is correct.

Student Guide. The large body of information which must be conveyed to the trainee suggested the need for its collection into book form. It was recognized that system strengths were in the area of providing demonstrations and practice situations, and that the presentation of voluminous textual materials on the CRT would not be the most efficient use of system resources. Besides this, the dot matrix CRT presentations can cause eye strain after prolonged exposure. It was therefore decided that a Student Guide should be provided to convey the bulk of the verbal information, and that the instruction presented on the system would reiterate only the salient aspects of a particular topic.

Realism. A compromise was effected between absolute fidelity to the operational environment and cost considerations based upon assumptions about the aspects of the stimulus which were most likely to facilitate transer of training. Thus, for the experimental prototype it was not deemed necessary to provide a look-alike radar console. However, it was considered to be important to provide a radar display which faithfully reproduced that in the operational environment.

TECHNICAL ASSUMPTIONS. The assumptions described below were made based upon the preliminary analysis of GCA-CTS requirements and the performance characteristics of the hardware and software components. It was assumed that:

a. Real-time speech recognition could be performed with sufficient accuracy to support the training concept by using commercially available hardware and augmented software;

b. A real-time radar simulation could be provided using commercially available hardware;

c. The speech synthesizer output would be intelligible to all users;

d. A speech digitizer could be designed to record and play back trainee speech (this assumption was necessitated by the fact that the audio disk unit which was originally proposed was not available at hardware procurement time);

e. The computational system resources (especially the single 10 Mbyte disk and the interprocessor bus transfer rates) would adequately support the processing requirements;

f. Commercially available support software (Data General's Real Time Disk Operating System (RDOS), Fortran 6, the load on call overlay manager and the graphics library package) would likewise support GCA-CTS processing with minimal modification.

TECHNOLOGY

The GCA-CTS was developed to evaluate the automated speech technologies as exemplified in the Threshold 500 voice input preprocessor and the Votrax VS-6.4 speech synthesizer. The application of these commercially available devices is not as easy as interfacing a CRT to a system, although at the hardware level there is little difference. For example, the speech recognition algorithms developed by the manufacturer work admirably when used in a standalone mode, over a relatively short vocabulary, and by an experienced user;

but in GCA-CTS they must work in parallel with many concurrent processes, over a long vocabulary, and by a naive user. Furthermore, user acceptance would be seriously impacted if the first experience with the system required the repetition of over 100 phrases ten times each! The challenge in this development effort has not been in the area of hardware integration, but rather in creating an environment in which this highly specialized hardware can be used effectively.

The record/playback technology used in the GCA-CTS also falls under the category of automated speech technology, and although it was not originally specified as an aspect of the technology which was to be evaluated, the exigencies of the project gave us the opportunity to do so. The devices are briefly described in the following paragraphs.

SPEECH RECOGNITION TECHNOLOGY. The Threshold 500, manufactured by Threshold Technology, Inc., was chosen as the device which exemplified the state of the speech recognition art (in 1977). It is an isolated phrase recognition system which operates by sampling the speech input every 2.2 msec and checking for the presence or absence of 30 features in the input stream. These features are of two types. About half of them are related to the relative energy content of specific spectral bands, and the rest result from logical and analog operations on the short-term power spectrum. Most of the latter features are attempts to detect phonemes or phoneme groups.

The information is transferred to the computer for storage as a bit pattern in which a bit is set if the feature was detected. When a pause of approximately 1/4 second is detected, the recognition algorithm time normalizes the collected data by forming two input feature patterns with 16 and 32 ime slots, respectively. (It should be noted that a double buffering scheme ised so that no data are lost if the user begins speaking immediately the 1/4 second pause. The time normalization and pattern recognition processes operate in parallel with subsequent speech data acquisition.) These input feature patterns are compared with previously collected reference patterns on a bit-by-bit basis. A three-pass search is conducted if necessary to find the reference pattern which most closely matches the input feature pattern. The algorithm then outputs the highest scoring pattern number as its recognition choice. If two scores are not significantly different, a second choice recognition is also supplied. Resolution is accomplished outside the recognition algorithm by semantic processing in the speech understanding subsystem.

This device is attached to the operating system interrupt structure at run time through the use of the .IDEF system call. The vendor-supplied device driver was modified extensively to accommodate the necessary double-buffering, and to work in the vectored-interrupt architecture of the Eclipse computer.

The primary risk areas in the successful incorporation of this technology are: the requirement to recognize a large vocabulary which includes many similar phrases and phrases which differ significantly in length; and the requirement to recognize this vocabulary when it is used by a naive user of the speech recognition equipment and who furthermore can be expected to change the way he speaks as his personal style evolves.

SPEECH SYNTHESIS TECHNOLOGY. The Votrax VS-6.4 speech synthesizer, manufactured by the Vocal Interface Division of Federal Screw Works, was chosen to be the GCA-CTS voice. The manufacturer describes the unit as an "electronic simulation of the human brain-vocal system." It electronically generates 63 phoneme-like sounds optimized for the Midwestern American English dialect. One of four levels of inflection is associated with each phoneme, so it is possible to formulate natural-sounding statements, questions, or even simple songs.

Unlike the Threshold 500, the device driver for the Votrax is built into the operating system. This enables the device to be used easily for vocabulary development as well as for speech output during GCA-CTS operation. The driver, developed by Logicon, recognizes two types of input: octal inflection/phoneme codes which it sends to the device directly; and ASCII phoneme names which it translates automatically to octal inflection/phoneme codes. This latter feature simplifies vocabulary development immensely because it enables user-oriented dialog with the device.

The primary risk area in the incorporation of this device is in the area of user acceptance of the speech quality. The device has a slightly artificial sound, yet most listeners report that it is highly intelligible.

SPEECH RECORD/PLAYBACK TECHNOLOGY. The requirement to automatically record and play back the trainee's speech was evaluated at proposal time and a random-access audio disk recorder with a Logicon-designed interface was specified. However, by hardware procurement time, this device was no longer available. The market was surveyed for a replacement device, but none was found to meet the functional requirements. For example, a computer-controlled audio tape recorder could not offer the precise control needed to satisfy the requirement for stopping and restarting the replay of the student speech after error explanations. The solution was to design and build a special purpose device based upon the newly developed continuously variable slope delta modulator integrated circuit technology. The chip encodes audio input at a rate of 16,000 bits per second. The decoding of these data produces an intelligible (though not high fidelity) replay of speech. This capability was used in several unique ways in the GCA-CTS. First, an audio recording of the trainee's speech is made during a problem and is replayed in synchrony with the aircraft dynamics at trainee request. This replay gives the trainee a basis for understanding his mistakes, especially those due to stylization errors. Secondly, the device is used to prompt the trainee during voice data collection. To ensure good speech recognition, it is extremely important to elicit natural sounding speech samples. This cannot necessarily be done by prompting on the CRT or speech synthesizer. In GCA-CTS, speech samples can be elicited in context, using the speech digitizer to prompt the trainee with his own voice. Finally, the system was designed to be capable of giving demonstrations with the student's voice as that of the final controller.

This data channel device is attached to the RDOS interrupt structure at run time. The primary risk factor in incorporating the device was its relatively high data rate. The GCA-CTS mass storage unit was chosen before the need to store the extensive courseware files was identified, and with the assumption that audio data would be stored on a separate medium. The need to save speech data for two 8-minute problems plus the digitized recording of the student's verbalizations of the GCA vocabulary items makes it impossible to store data for more than one trainee per removable cartridge disk. Tne high data rate also adds additional disk I/O overhead.

SECTION II

METHOD

This section discusses the GCA-CTS project from an historical perspective, explaining who did the work and how it was accomplished.

ORGANIZATION AND STAFFING

The GCA-CTS project was undertaken by the Advanced Systems Department, Tactical and Training Systems Division of Logicon, Inc., whose charter it is to apply emerging technologies to solve the identified problems of customers within the systems engineering research and development environment. The Advanced Systems Group is not production oriented; rather it is a small group of individuals with multidisciplinary skills. Most of the work done in the department involves designing and building new systems which employ new technologies in unique ways. Because of the nature of this work, very often contract specifications are like the GCA-CTS specification: rather general in focus. An important part of the task involves the definition of what needs to be done, precisely because the problem has not yet been solved.

Project responsibility centers with the Technical Contract Manager who is responsible for identifying, allocating and monitoring the resources needed to get the job done, and in general, for the administrative side of the project. Working closely with the Technical Contract Manager is the Project Leader who has responsibility for the technical side of the effort. The Project Leader coordinates the efforts of the project team members. For the GCA-CTS project, these team members included a training psychologist, instructional system designers, a hardware design engineer, hardware engineers, mathematical analysts, and software specialists.

ACTIVITY

This section describes this 22,000 labor hour development effort (see Table 2), highlighting the problems to be solved and their solutions.

TRAINING FUNCTIONAL DESIGN REPORT PREPARATION. This baseline document described the behavioral objectives of PAR controller training, defined a syllabus through which this training was to occur, and provided a functional specification for the system.

Behavioral Objectives Report. The first major section of the Training/Functional Design Report (Hicklin, Nowell and Petersen, 1978) consisted of a behavioral objectives report. This was prepared by reviewing Government documents, especially Training Characteristics of the Ground Controlled Approach Radar Controller Training System (Breaux, 1976), and holding discussions with NATTC personnel in December, 1977. The analysis of the GCA

TABLE 2. MAN-HOURS EXPENDED (ROUNDED) FOR GCA-CTS DEVELOPMENT

Activity	Approximate Labor Hours
Management, Integrated Logistics Support Planning, Quality Control	1020
Functional Definition	1540
Software Design	2700
Implementation	9770
Hardware Engineering	2190
Delivery and Support	1710
Publications Support	2270

final approach controller's responsibilities included the following major GCA task areas:

a. Equipment setup and usage

b. Pattern controller coordination

c. Precision approach control

d. Approach termination procedures

e. Emergency procedures.

This section of the report was prepared using the guidelines laid down by the contract. Briefly, as an organizational and analytical aid in dissecting the behavioral aspects of performing this task, a hierarchy of behaviors was developed. Four mission objectives were derived from the course objective. Each mission objective describes distinct segments of the GCA controller's task. The attainment of a mission objective requires that several complex behaviors be performed at appropriate times. Mastery of each of these complex behaviors is described as a terminal objective within the larger mission objective context. Finally, the complex behaviors can be subdivided into their constituent simple or enabling behaviors. Figure 1 illustrates this form of hierarchy.

The distinction among levels of objectives consists in the following: The enabling behaviors are independent and directly measurable. Strict standards can be defined for the performance of each enabling behavior. At the terminal objective level, the relative significance of violations of these standards can be taken into account. Finally, at the mission objective level, rules for combining the behaviors can be applied. The course objective stands

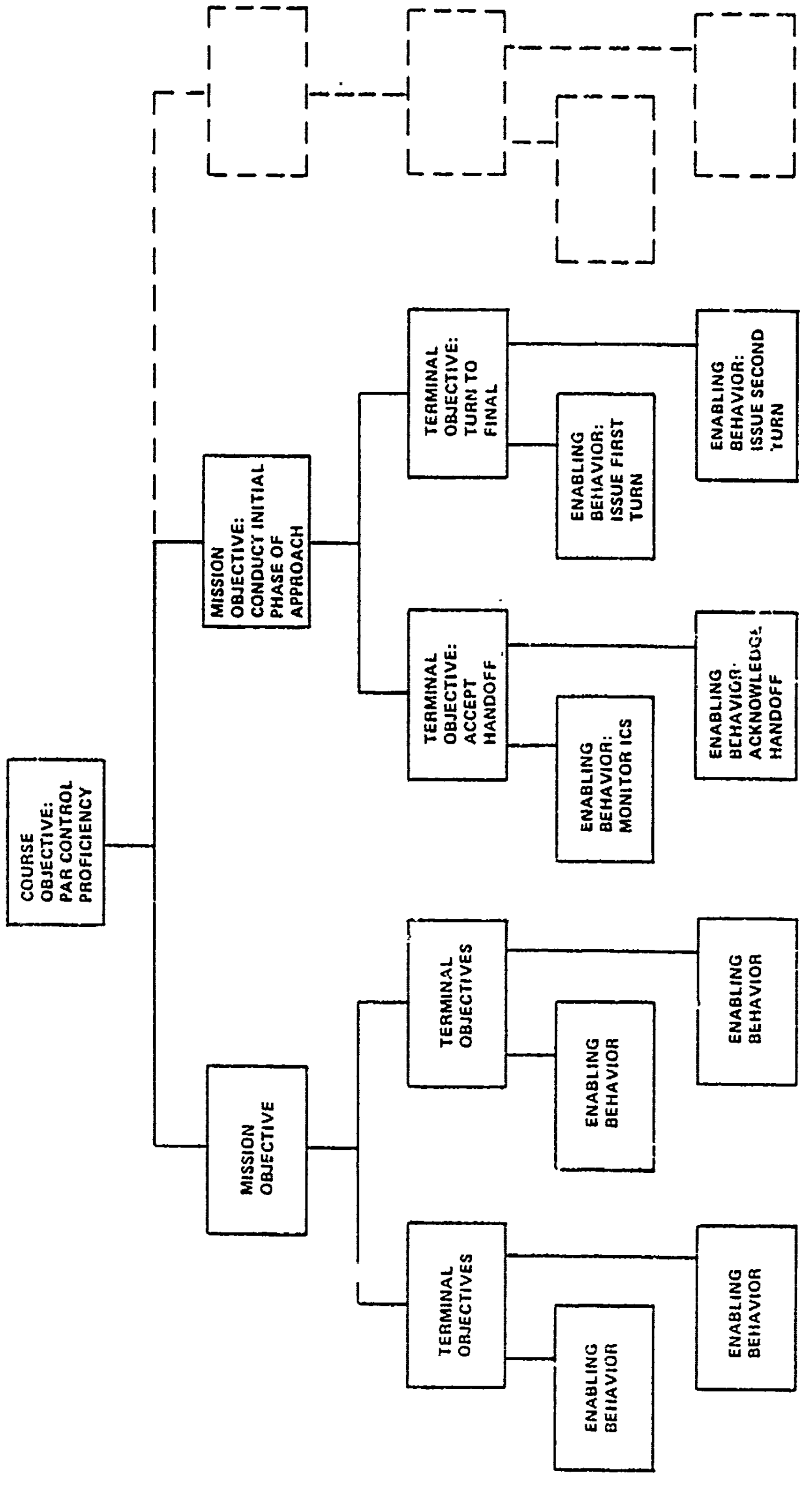

Figure 1. Formal Structure of the Behavioral Objectives Hierarchy and Example

as the final cause of all the objectives and serves therefore as a system design goal and as the measure of system performance.

Syllabus. The second major section of the Training/Functional Design Report (Hicklin, Nowell and Petersen, 1978) consisted in a preliminary course syllabus which showed the order of presentation of the topics defined in the behavioral objectives section. Although the syllabus underwent some refinement during the course of the project, the organizational principles laid down in this report guided all modifications.

The structural elements of the course syllabus are the levels of achievement shown in Figure 2. The illustration was drawn from the Student Guide (Hicklin et al., 1980b) and emphasizes the point that the trainee attains to control proficiency through the attainment of intermediate goals. The levels themselves are organized in such a way that the trainee's previously acquired skills serve as a foundation for the new material. Thus, for example, the first task the trainee learns is azimuth control, building upon previously acquired surveillance radar skills. Glidepath control procedures then build upon target division skills acquired in the level devoted to azimuth control, and so on.

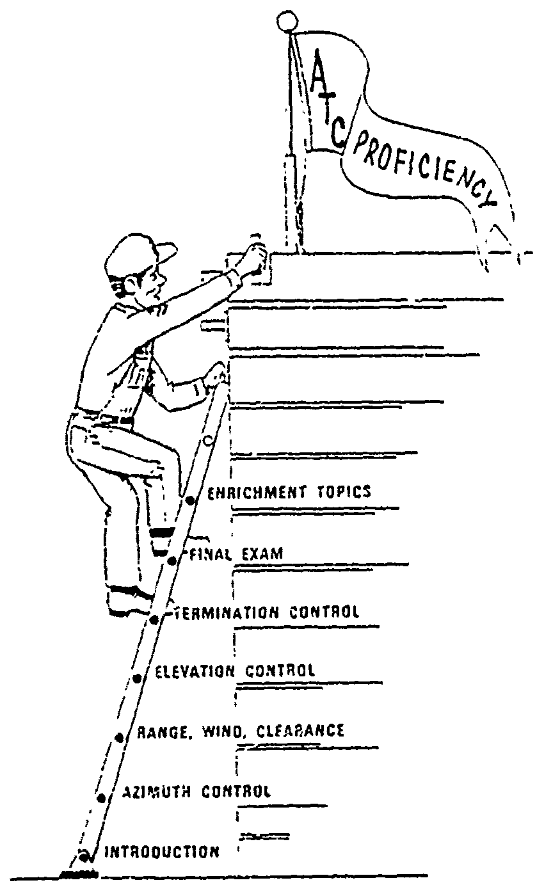

Figure 2. GCA-CTS Levels of Achievement

Within each level of achievement, a phased approach is used to instruct the trainee and to let him or her practice and therefore acquire the new skill. In general, a small, self-contained aspect of the task is explained in an interactive teaching mode and the system collects the voice reference patterns needed to recognize the trainee's voice as shown in Figure 3. An optional commented practice phase is offered as shown in Figure 4 in which the system provides immediate feedback on the trainee's performance on the new material. Finally, a graded practice phase is provided in which realistic control problems are given as shown in Figure 5. The trainee practices the new skill in the simulated environment and has an opportunity to integrate the new material with previously acquired skills. A final examination in the form of a performance test which is similar to graded practice, is also provided.

Functional Specification. The third major section of the Training/Functional Design Report (Hicklin, Nowell and Petersen, 1978) was the specification which described the major functions required to provide these training capabilities. It detailed special simulation requirements as well as the functional organization of software modules. It detailed the precise vocabulary to be recognized by the system, as well as the pilot and pattern controller dialog to be synthesized by the system. Supporting material included a cross-reference table showing the relation between the syllabus tasks and the behavioral objectives.

Problem Areas. The functional definition activities pointed to several areas which had not been implemented in the laboratory system; and, based upon the Statement of Work, were not fully appreciated in scoping the experimental prototype system. These included:

a. The inclusion of course messages and course trends in the repertoire of transmissions used throughout the approach.

b. Extensive communications with the pattern controller.

c. The need to simulate and teach radar servo control.

d. The significance of a fairly extensive Student Guide to support the highly automated, total training system concept.

e. The extent to which the proper use of equipment must be taught by the GCA-CTS.

f. The importance of providing remediation in the training program.

g. The need to provide commented practice.

In addition, the need to design and build a separate functional input panel at the trainee station was confirmed. (A contract modification was negotiated to incorporate the required enhancements.)

SYSTEM CONFIGURATION/FACILITIES REPORT PREPARATION. This report was prepared to document the results of the hardware and software design effort. Highlights from this effort are discussed below.

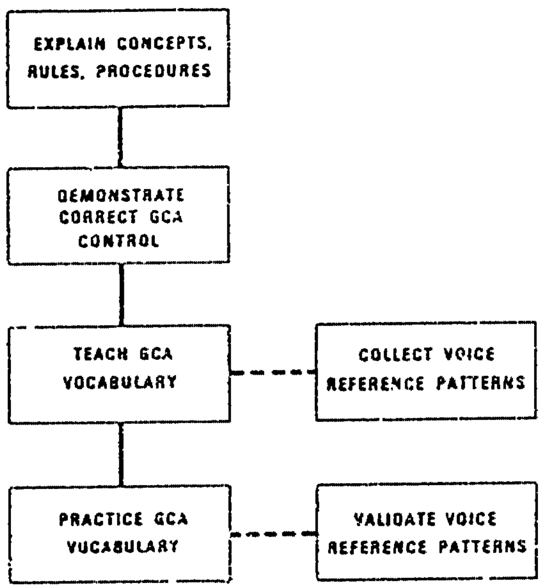

Figure 3. Phase 1: Interactive Teaching and Voice Data Collection

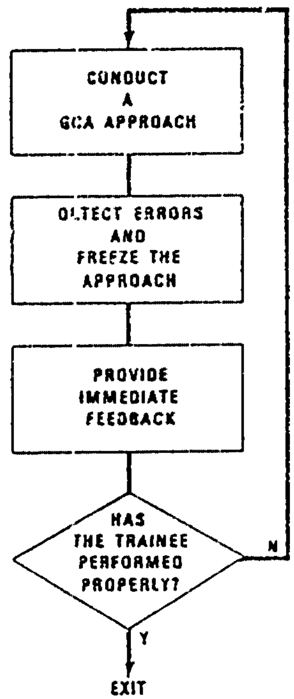

Figure 4. Phase 2: Commented Practice ("Freeze and Feedback")

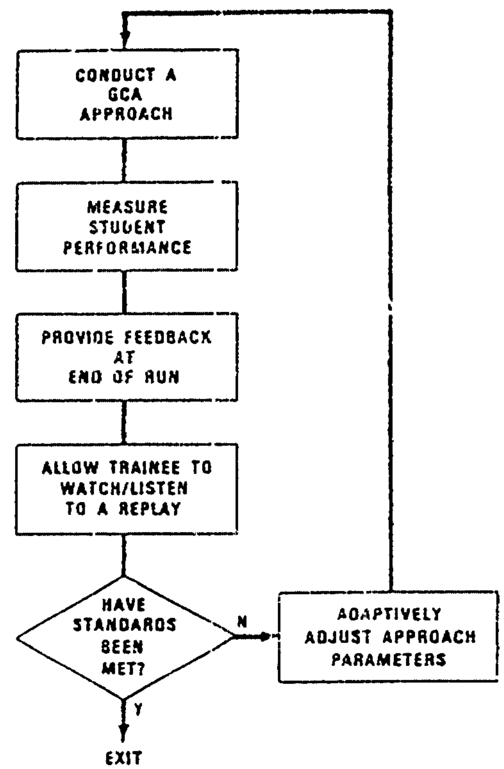

Figure 5. Phase 3: Graded Practice

Hardware and Software Environment. Commercial hardware and software were used to the fullest extent possible in the GCA-CTS. The computational system consisted of two Data General S/130 minicomputers and associated peripherals supplied as Government-furnished equipment. Logicon procured the additional peripheral equipment including a Tally 1602 printer, Megatek MG552 graphics display terminal with joystick and an associated software package, a Votrax speech synthesizer and a Threshold 500 voice input preprocessor.

Data General's Fortran 5 language with its interface to the multitask scheduler in their Real-Time Disk Operating System was selected as the primary implementation language. Data General's Macro Assembly language was selected for use in those cases where a high-level implementation was impossible, as for example in the coding of device drivers. The other commercial software that was employed was the Megatek graphics package. This package actually required extensive modification since it was written in Fortran IV-compatible assembly language for use in an unmapped, single-task environment.

Special Purpose Hardware Design. The unique functional requirements of the GCA-CTS necessitated the design of some special purpose hardware. This hardware included:

a. A record/playback device based upon the commercially available continuously variable slope delta modulator integrated circuit which digitally encodes audio input or decodes digital data to produce intelligible audio output at a rate of 16,000 bits per second.

b. A trainee panel incorporating simulated communications equipment, the simulated radar servo control, the speech preprocessor voice level meter and level control knob, and microphone and headset controls, and an attached simulated microphone footkey.

c. An instructor panel incorporating an intercom for monitoring and communicating with the trainee.

d. A junction panel for the distribution of outputs and acquisition of inputs from the special purpose hardware elements.

With respect to the design of the trainee panel, a question arose regarding the need to mount the joystick in a separate unit in order to more closely approximate the position of the servo control on the operational gear. The rationale for the hardware design is that as long as the control is mounted firmly in the same orientation as on the operational gear, is within easy reach on the same side of the display, and most importantly that its functional behavior simulates the actual gear very closely, then transfer of training should occur. The fact that the device was not be mounted on the simulated radar unit itself is not predicted to have a negative impact on transfer of training.

Although an audio disk unit had been proposed to satisfy the GCA-CTS record/playback requirements, the device was no longer available at the hardware procurement stage of the contract. The various recording devices on the market were again studied at this stage but none were found which could satisfy GCA-CTS requirements. It was for this reason that the special purpose speech digitizer was developed.

Software Design. A top-down approach was taken in the design of the GCA-CTS. The system inputs and outputs were defined, then a system hierarchy was designed, the major system elements within the topmost layer of the hierarchy were identified, and the data flow between them was defined. This was done for the other levels of the hierarchy in turn. These major system elements were then broken into their component parts, again beginning with those highest in the hierarchy and the process continued until individual software modules were identified with their inputs and outputs, and the data flow throughout the system had been defined. At this point, the design of the individual modules commenced.

More specifically, the design began with the traditional postulating of user inputs and system outputs. However, a problem arose at once as to how to conceive of GCA-CTS. Did each phase of each task require a separately compiled set of programs? Should GCA-CTS be conceived as four relatively independent systems on the basis of a temporal differentiation between phases 1, 2, 3 and replay? Could it be designed as one system in which the phase distinction was subsumed under training control? Since many system functions are used during more than one of the phases, as shown in Table 3, the last approach seemed best. Given this choice, it was necessary to devise a means whereby a particular phase control subprogram could orchestrate the system resources to provide information presentation, data collection, etc., as needed for the particular task. A design goal was to make each of these phase control subprograms as general as possible so that one phase 1 control subprogram handles phase 1 training for all tasks in the syllabus, and likewise one control subprogram handles each of the other phases and replay. A higher level executive was designed to invoke these control subprograms as necessary.

The syllabus was defined to be a file which consists of an ordered list of file names. There is one file name for every phase of every task in the syllabus. Associated with the file name is an indicator of the phase of instruction represented in the file. The training system executive was designed to retrieve the file names sequentially (when progress to the next sequential phase is appropriate), use the indicator to select the proper phase control subprogram and transfer control to it. The phase control subprogram then accesses the given file to retrieve run specific parameters or in the case of phase 1, the sequence of instruction.

As the next step, the inputs to the phase executives which produce the outputs required for training were elaborated so that the table-driven executives could be designed in accordance with the inputs specified for them. The concept of table-driven training makes modification of the training course straightforward and is probably the only way the ambitious GCA-CTS training requirements could have been met in a timely and cost effective manner.

After the training executive was designed, design work focused on phase 1 partly because of its extreme importance to the quality of training, and partly because this work directly impacted the design of the training materials

TABLE 3. FUNCTIONAL ELEMENTS OF THE GCA-CTS MODES OF OPERATION

Applicable Functions	Mode				
	Demonstration	Phase 1	Phase 2	Phase 3, P-run	Replay
Voice data collection		X			
Speech recognition			X	X	
Speech understanding			X	X	
Aircraft, pilot, environment	X	X	X	X	
Radar	X	X	X	X	
Display	X	X	X	X	X
Model controllers	X	X	X	X	
Performance measurement			X	X	
Keyboard input processing	X	X	X	X	X
IPB I/O processing	X	X	X	X	X
Trainee panel input processing		X	X	X	
Trainee panel output processing	X	X	X	X	X
Votrax output processing	X	X	X	X	X
Speech digitizer input processing		X		X	
Speech digitizer output processing	X	X	X	X	X
User Clocks	X	X	X	X	X

which was to commence early in the project. Phases 2 and 3 were not regarded as being as critical because they required fairly well understood simulations and their operation was similar in principle to the laboratory GCA-CTS. Phase 1, on the other hand, bears little resemblance to its earlier counterpart, the stand-alone voice data collection program. It provides great flexibility in the presentation of training materials, but, of course, has limitations which had to be laid out before the instructional technologist could design the course. A subject mattter expert and the system designers worked together to ensure the phase 1 executive would provide needed capabilities, and that costly, unneeded capabilities would be avoided.

After the design of the phase 1 executive was well underway, concern shifted to the other phase executives and the replay executive, and finally to their constituent modules. All of this work is thoroughly described in the report (Barber et al., 1978) and need not be detailed here. However, some of the salient features of the design which might otherwise be lost amidst the volume of detail are discussed briefly.

Table-Driven Design. The philosophy of a table-driven system of instruction is fundamental to GCA-CTS operation. Figure 6 illustrates the concept. To the left is a portion of the GCA-CTS syllabus. It is composed of comments and actual file names with associated phase specification information. The training control executive processes these file names sequentially (unless the adaptive scheduler or instructor directs it to otherwise). The processing merely involves calling upon the specified phase executive which, in turn, looks at the content of the specified file to obtain the sequence of instructions or the type of practice which is to be given. The courseware is thus independent of the training system software and can therefore be modified without the necessity of modifying the training system itself.

Although there were insufficient resources in this experimental prototype contract to do so, it should be noted that a courseware development language could easily be devised to generate the courseware tables through a user-oriented dialog.

Training Control. The training control executive is the primary GCA-CTS executive. It has the overall responsibility of providing the following capabilities:

a. Task selection and mode sequencing

b. Adaptive problem specification

c. Remedial problem selection

d. Feedback presentation

e. Performance test administration

f. Record keeping

g. Report preparation

h. Special request processing

i. Demonstration mode presentation

Interesting elements of the design are those related to adaptive training. Although the basic syllabus is rigidly defined in terms of the order of topic presentation, as it must be to accommodate the exigencies of the "errorless" learning philosophy, the automated instructor can adapt the course of training in terms of problem difficulty, number of practice problems given, and remediation as follows. After every problem, the system evaluates

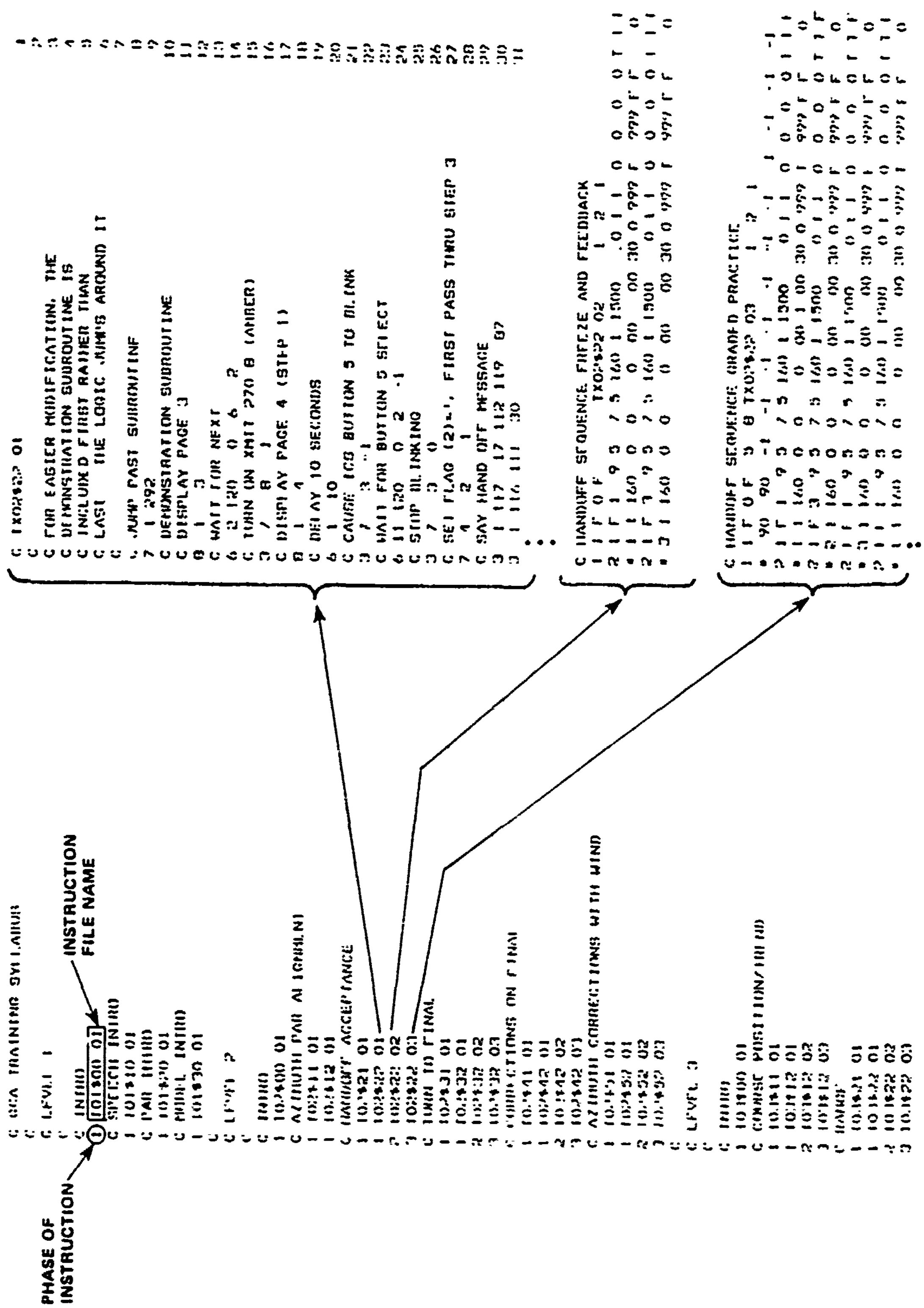

Figure 6. Table-Driven Design

the trainee's performance on the previously learned material. If this performance is significantly less than the average performance attained in the previous task, it is assumed that the trainee is having trouble integrating the new material with the old. GCA-CTS then adjusts the difficulty of the upcoming problem in such a way that the skill he or she is having trouble with will be easier to apply. For each of the skill categories shown in Table 4 the system asks: Is the student's performance significantly worse than it was on the average in the last phase 3 task? If the answer is yes, then the adjustments shown in the table are made.

TABLE 4. ADAPTIVE PROBLEM SELECTION

Skill Category	Adaptation Applied
Heading transmissions	Set wind variability, correlation time and gusting to easiest values.
Azimuth position and trend	Select best pilot, slowest aircraft
Glidepath position and trend	Select best pilot, slowest aircraft
Range calls	Select slowest aircraft

When the minimum number of approaches specified for the task have been completed, GCA-CTS will ask whether the student's performance meets the criteria established for the new material at this level. If not, another problem is given. This continues until the student's performance reaches the criteria, or until he or she has completed the maximum number of problems for the task.

When the trainee has completed the requirements for a particular task, the system either advances him or her to the next task, or selects an appropriate remedial exercise. Remediation is chosen if the trainee's average score on the previously learned material for this task does not meet the established criteria. The remedial problem or problems selected depend upon the skill category for which the low score was obtained, and upon the level which has been reached in the syllabus. The time constraints imposed upon the present course caused the remedial problems originally specified for several levels in the syllabus to be deleted. However, the potential exists to provide a variety of remedial learning experiences.

The system is also adaptive to the trainee's expressed needs. First, commented practice (freeze and feedback) problems are optional. Secondly, replay with or without error reporting is optional. Thirdly, the voice testing and voice data collection modes are available on request.

Feedback is automatically given to the trainee after every problem in the form of a performance summary (Figure 7) and optional replay. When the trainee has attained to proficiency, a performance test or final examination is automatically given and a report is prepared for the learning supervisor as shown in Figure 8. The system was designed to keep records about trainee performance and to prepare hard copies of performance reports at instructor

PERFORMANCE FEEDBACK

You have completed 3 of the problems in task: T04$32.03
You must complete a minimum of 5 problems but not more than 10 problems.
Your performance on new material:

Glidepath, position and trend	Needs work

Your performance on other tasks:

Accepting handoff	Perfect
Radio check	Perfect
Turn-to-final	Perfect
Approaching glidepath	Satisfactory
Heading transmissions	Perfect
Azimuth position and trend	Satisfactory
Range calls	Perfect
Clearance	Needs work
Handoff and rollout	Satisfactory
Transmission break	Perfect

Figure 7. Sample Feedback Provided at the Trainee Station CRT After Every Graded Practice Problem

request. These performance reports were designed to provide varying levels of detail. The first offers verbal generalizations about task performance (Figure 9), a second provides scores attained for each problem in a task (Figure 10) and a third shows specific information about the errors made on a particular task (Figure 11).

The preparation of these reports is handled by the special request processing feature of the training control executive. Tables 5 and 6 show these and other options available to the instructor and trainee.

A demonstration mode was also designed for use both in the instructional phase and by the training control executive. Its purpose in the instructional phase is to show the trainee how to perform specific procedures. It is also initiated by the training control mode on system startup and whenever no trainee is signed on to the system. The purpose of the latter use is to provide a realistic environment for assuming radar control duties. The student is taught that he or she is responsible for checking the alignment of the PAR radars, as is the case in the operational environment. The student is taught to check this as soon as possible after taking over the operator position. The procedure involves ensuring that no other controller is using the gear, then observing the radar return from fixed reflectors with respect to the electronically generated cursors. To provide the appropriate stimulus for eliciting this response, it was necessary to have the system take the role of another controller and conduct approaches. Thus, even though the system provides many different kinds of instruction, it always simulates an operational environment during the time the trainee is taking over the control position.

PERFORMANCE RUN SUMMARY REPORT

NAME: HICKLIN

DATE: 8-16-1979 TIME: 1855

REC	SPKR	RANGE	TIME	ADVISORY
8	PTN	9.5	0	Position 4 Hand off Right base
9	PTN	9.5	0	Marine 687 A6 Low approach Button 1
19	GCA	9.5	1	Position 4 roger.
***			11	REFERENCE # 1 Handoff not acknowledged within 10 secs of issuance
21	PTN	9.5	15	Marine 687 After completing Low approach Climb and maintain 1500
22	PTN	9.5	15	Turn right heading 2 7 0
23	PTN	9.5	15	Over
26	PLT	9.5	27	Roger
30	PTN	9.5	28	Marine 687 Turn right heading 150
31	PTN	9.5	28	Over
34	PLT	9.5	35	Roger
45	GCA	8.6	68	Marine 687 Radar button 1.
***			78	REFERENCE # 2 Proper frequency not selected for radio contact

Figure 8. Sample P-Run Report

```
NAME: HARMON                    WILBUR           DATE: 3-13-1980   TIME:1958

                          PERFORMANCE ANALYSIS:
 STRENGTHS                  BORDERLINE                WEAKNESSES

 HANDOFF                    GLIDEPATH POSITION/TREND  TURN TO FINAL
 RADIO CHECK                                          HEADING TRANSMISSIONS
 APPROACHING GLIDEPATH                                RANGE CALLS
 AZIMUTH POSITION/TREND                               EMERGENCY WAVEOFFS
 DECISION HEIGHT MESSAGE
 CLEARANCE REQUESTS
 LANDING THRESHOLD
 ROLLOUT OR  HANDOFF
 TRANSMISSION BREAK
 TRANSMISSION RATE

 STUDENT WAS ADVANCED TO PRESENT LEVEL AFTER COMPLETING  1 RUNS
 NO REMEDIATION NEEDED

 TOTAL SYSTEM TIME TO DATE:     16 HOURS AND    17 MINUTES
```

Figure 9. Sample Task Summary Report

Phase 1 Training Executive. The phase 1 executive interprets a courseware presentation language to provide multimedia computer-aided instruction and transparent voice data collection. Table 7 shows the types of instructions which are interpreted by this executive.

Phase 2 and Phase 3 Training Executives. In contrast to the phase 1 executive, the routines controlling the operation of the other phases of training are relatively simple. They are only required to initialize the conditions specified for a given problem, start the simulation executives, then wait for problem termination and provide feedback.

Although the simulations are dynamic and responsive to the trainee's performance, these executives give the courseware designer specific control over the initial conditions of the simulation parameters shown in Table 8. Both the phase 2 and phase 3 executives are capable of providing practice problems specified precisely in terms of the parameters shown in the table. While this degree of control is necessary, it was obvious that the specification in this format of ten or twenty similar but unique practice problems for each task in the syllabus would be laborious for the courseware designer. Therefore, a multipossibility problem specification format was also designed. The courseware designer can simply specify the acceptable range of each of the problem-specific parameters, and the phase 3 executive selects individual run parameters randomly from ones in that range.

Replay. In the GCA-CTS, a great deal of attention was directed to the design of feedback which would enable the user to understand and learn from his mistakes. In this unique training environment, these mistakes include stylization errors which cause recognition failures. A replay capability was added

NAME: HARMON WILBUR PERFORMANCE ON TASK: 103432.03

DATE: 3-6 1980 TIME: 1442

RUN#	HO*	RCb	TTF	AGP	HA	APT	GPT	RCa	DH	CL	OLT	HR	NG	HC	EW	LAA	TB	TR
27	100	90	100		100	100		80		90								
28	100	80	[illegible]		100	50		76		95								
29	100	100	100		90	80		60		50								
30	100	100	100		99	80		80		100								
31	100	90	64		100	100		67		100								

*Explanations of abbreviations:
HO Handoff; RCb=Radio check; TTF=turn to final; AGP=Approaching glidepath; HA=Heading transmissions; APT=Azimuth position and trend; GPT=Glidepath position and trend; RCa=Range calls; DH=Decision height; CL=Clearance; OLT=Over landing threshold; HR=Handoff or rollout; NG=No gyro; HC=Heading corrections; EW=Emergency waveoff; LAA=Low altitude alert; TB=transmission break; TR=trans

Figure 10. Sample Task Performance Report

NAME: HARMON WILBUR PERFORMANCE ON PROBLEM 30, TASK: T03$32.03

DATE: 1 6 1980 TIME: 1442

SUBJECT AREA	SCORE	ERRORS
ACCEPTING HANDOFF	100	NONE
RADIO CHECK	100	NONE
TURN-TO FINAL	100	NONE
HEADING TRANSMISSIONS	99	You gave a turn of 1 degree
AZIMUTH POSITION AND TREND	98	Incorrect position given in azimuth position call
RANGE CALLS	80	Range call omitted Incorrect range used in range call
CLEARANCE	100	NONE

Figure 11. Sample Problem Performance Report

TABLE 5. FUNCTIONS OF KEYS AT INSTRUCTOR STATION

Key Name	Function
MENU	Displays on the CRT the legal keys for the current situation.
NEW T/E	Initializes new trainee files.
INIT VOICE TEST	Causes the system to enter the speech validation mode at the conclusion of the present exercise. In the validation mode, the system will attempt to echo the spoken phrase.
STOP VOICE TEST	Terminates speech validation.
YES	Used for responses to queries.
NO	Used for responses to queries.
STATS	Displays student status information on the CRT.
PRINT STAT	Provides detailed hard copy status reports.
↑STOP	Causes the GCA-CTS program to terminate. Both processors return to the CLI.
WAIT	Temporarily stops or freezes a demo or phase 3 run.
CONT	Continues a run suspended by a WAIT, continues training after ABORT.
ABORT	Stops the current run.
OVERRIDE	Allows the instructor to override GCA-CTS' problem selection.
INIT NEW R/T	Causes the speech data collection mode to be started after the completion of the present run.
↑MENU	A debug option. By default, CTRL C is disabled. Pressing these keys enables it. A subsequent press again disables CTRL C.
REPLA	Causes replay of student's performance run after completion of the present run.
MOD	This key invokes the replay file editor which corrects any mis-recognitions in the replay file. Training is suspended during this operation.
INIT T/E KBRD	Activates the instructor functions on the trainee keyboard.
EXIT T/E KBRD	Deactivates instructor functions on the trainee keyboard.

TABLE 6. FUNCTIONS OF KEYS AT TRAINEE STATION

Key Name	Function
MENU	Displays on the CRT the legal keys for the current situation.
HELP	Displays a request for assistance on the instructor console.
INIT VOICE TEST	Causes the system to enter the speech validation mode at the conclusion of the present exercise. In the validation mode, the system will attempt to echo the spoken phrase.
STOP VOICE TEST	Terminates speech validation.
ALIGN	Sets centerline range and touchdown reflectors into proper alignment.
NEXT	Continues with the next frame of the lesson.
YES	Used for responses to queries.
NO	Used for responses to queries.
HELLO	Initiates student sign-on procedure.
BYE	Terminates the session at the completion of the current problem. Demo will be started.
WAIT	Temporarily stops a demo or phase 3 run.
CONT	Continues a run suspended by a WAIT, continues training after ABORT.
ABORT	Stops the current run.
OVERRIDE	Allows the instructor to override GCA-CTS' problem selection.
INIT NEW R/T	Causes the speech data collection mode to be started after the completion of the present run.
[illegible] MENU	Toggles CTRL C enable on and off.
EXIT T/E KBRD	Deactivates instructor functions on the trainee keyboard.

TABLE 7. INSTRUCTIONS INTERPRETED BY THE PHASE 1 EXECUTIVE

Instruction Type	Instructions Interpreted
Voice Data Collection (VDC)	Start VDC.
	Collect a speech sample of phrase(s).
	Form voice reference pattern(s) for phrase(s).
	Validate specified phrase(s) to a specified percentage accuracy.
	Perform validation without prompting: echo whatever is spoken.
	Stop VDC.
Display	Initialize display list.
	Start display processor.
	Turn a picture on.*
	Display aircraft position update.
	Display wind information update.
	Display text message.
	Fade target trails.
	Display long trails.
	Turn a picture off.
	Turn all pictures off.
	Stop display processor.
Prompts	Output phrase(s) to the speech synthesizer.
	Output phrase(s) to the trainee's CRT.
	Output phrase(s) to the speech digitizer.
	Activate model controller, using the specified device for its output.
	Terminate model controller and demonstration activity.
	Store digitized input for a given phrase (prompting generated automatically).
	Cause the specified change to the trainee panel display.
Aircraft simulation	Initialize aircraft/pilot/environment (APE) model to the specified conditions.
	Initiate aircraft dynamics.
	Freeze aircraft dynamics.
	Terminate APE.

---

*The separate elements of the PAR display are referred to as "pictures." Each can be displayed in isolation. These separate elements are: Azimuth cursor and outline, azimuth hashmarks, azimuth target, azimuth trail, azimuth long trail, reflectors on azimuth display, elevation cursor and outline, elevation hashmarks, elevation target, elevation trail, elevation long trail, reflectors on elevation display, wind information, and text.

TABLE 7. INSTRUCTIONS INTERPRETED BY THE PHASE 1 EXECUTIVE (CONT)

Instruction Type	Instructions Interpreted
Radar simulation	Activate elevation radar servo, initializing position of servo and/or of centerline reflector to the specified conditions. Activate elevation radar servo, initializing position of servo and/or of range alignment to the specified conditions. Turn off azimuth servo, reset alignment. Activate azimuth radar servo, initializing position of servo and/or of touchdown relector to the specified conditions. Turn off elevation servo, reset alignment. Servo radar(s) to specified position.
Wait conditions	Delay x seconds. Wait for keyboard entry of special key(s) (e.g., YES, NO, etc.). Skip the specified number of instructions on timeout or for each input option. Wait for keyboard entry of standard key(s). Skip the specified number of instructions on timeout or for each input option. Wait for trainee to servo azimuth antenna to a specified zone; skip the specified number of instructions on timeout. Wait for trainee to servo elevation antenna to a specified zone; skip the specified number of instructions on timeout. Wait for the aircraft to enter an azimuth position zone. Wait for the aircraft to enter an elevation position zone. Wait for the aircraft to reach a given range. Wait for the speech synthesizer to finish speaking; skip instructions on timeout. Wait for the end of a digitized speech utterance; skip instructions on timeout. Wait for the specified change in trainee panel status. Wait for end of trainee voice input; skip the specified number of instructions on timeout.
Sequence instructions	Skip a specified number of instructions. Jump to a subroutine. Up to five abnormal returns can be specified. Subroutines can be nested to five levels.

TABLE 7. INSTRUCTIONS INTERPRETED BY THE PHASE 1 EXECUTIVE (CONT)

Instruction Type	Instructions Interpreted
Sequence instructions (Cont)	Return from subroutine. Set a flag to a specified condition. Skip instructions based upon flag condition. Return to the start of the instruction file.
Text presentation	Display a specified logical page from the text file. Display a specified message on the trainee's CRT. Display a specified message on the learning supervisor's CRT.

to the laboratory GCA-CTS and it was found to be an effective feedback technique for procedural errors. In that system the speech synthesizer repeats the trainee's advisories and gives rule explanations when an error is encountered. The investigator noted, however, that "message not understood" reports proved especially frustrating to the students. Many times a student was convinced that he had uttered the correct advisory but he had no way to argue with the computer or to understand why the recognition failure occurred. For example, he would remember that he had said "slightly above glidepath" but would not realize he had paused in mid-phrase, making the advisory unintelligible to speech recognition.

To add to replay's usefulness as a feedback technique and to reduce the frustrations associated with recognition failures, a speech input digitizer was designed to record the trainee's utterances. The replay consists of an actual recording of the trainee's speech, synchronized with the visual display. From this, the student should be able to understand the cause of any recognition failures which occur. In addition to enhancing the student's acceptance of training system decisions, the replay provides the instructor with an excellent tool. After the run the student and instructor can review the run together. This provides a forum for discussion of such things as subtle points of style. At the trainee's request, replay will also pause and explain the errors which were detected by the performance measurement subsystem.

There are two ways in which the design of a replay capability can be approached. First, the "replay" can in reality be a recreation of the original dynamics. The alternative is to sample and save data during the original approach, then use these actual run data to provide a replay. The latter approach provides a true replay and so is aesthetically more pleasing. It also proved to be the more practical approach in this application for the following reasons.

First, the observable results (i.e., those required for replay) of the complex aircraft/pilot/environment and radar simulations are described by the eight words of target position and wind information transferred to the display processor every half second (the sweep rate). This data rate is miniscule compared with the data rate required to encode the trainee's speech (1000

TABLE 8. COURSEWARE SPECIFIED PROBLEM PARAMETERS

Task Specific Parameters

Azimuth radar display: off, on, on without hashmarks
Azimuth servo: off or on
Elevation radar display: off, on, on without hashmarks
Elevation servo: off or on
Minimum number of runs
Maximum number of runs
Text file name
Performance measurement variables relating to this task
Performance Criteria for advancement to next task

Problem Specific Parameters

Aircraft type: U-21, A6, P3, T38
Starting range from touchdown
Ending range
Initial aircraft altitude specified in feet or by elevation position zone
Initial aircraft offset from the centerline specified in miles or by azimuth position zone
Pilot type: 1 (best) - 5 (worst)
Type of flight:

1. Pilot responds normally to control instructions
2. Restrict aircraft position to the specified contiguous azimuth zones.
3. Restrict aircraft position to the specified contiguous elevation zones.
4. Restrict aircraft position to the specified contiguous azimuth and elevation zones.

Handoff: given or not given
Azimuth target display: on or off
Elevation target display: on or off

Approach type: Full stop, low approach, touch-and-go, short approach, no-gyro approach.
Clearance: Clearance given at first request, continue then clear at 2 miles, not given, wave off, clearance cancelled.
Wind information: Mean heading, mean wind speed, mean gust speed, mean gust duration, fraction of time gusting occurs, wind variability, wind speed correlation time.

Ceiling height
Wheels check: Pilot does or does not respond to radio check with "...wheels down..."
Number of seconds pilot waits before assuming lost communications.
Low altitude alert: if specified, force aircraft to descend to a point which requires a low altitude alert be given.
Gyro failure: if specified, disable the gyro compass at the specified range.

words per second) and so data acquisition does not pose a significant additional burden, particularly compared to the processing burden required to generate them.

Secondly, data about the other observable aspects of the problem (e.g., pattern controller dialog, pilot responses, trainee button presses resulting in light display changes, servo manipulation, etc.) must be saved anyway for performance measurement and for the performance test summary report, therefore, no additional acquisition of data is required to perform replay in this way.

Finally, as a matter of practical concern, even if replay had not been designed in the way it was, similar file structures would have to have been devised as debugging tools to provide a sampling of event-driven as well as cycle-by-cycle system outputs for analysis. Data saving for replay thus alleviated the need to design an additional debugging tool. This design philosophy led to the design of three replay data structures: a digitized speech data file, a display data file, and an activity or performance data file.

Aspects of the design of interest from an instructional perspective include the provision of error reports, rule explanations, and additional information to the trainee during annotated replay. These reports are varied automatically. In general, for each error the system detects, the system provides information about the basis for the error report ("you were understood to say..."), one of two possible rule explanations, and may randomly append a third statement which amplifies the rule or explains the consequences of failing to abide by it. The correct transmission or state-of-the-world information is provided to help the trainee understand the correct procedure.

A replay without error reporting capability was also designed in this experimental prototype system so that replays could be observed and scored by independent observers for comparison to the GCA-CTS scoring algorithm.

Performance Measurement Subsystem. In the GCA-CTS, performance data are required to provide the following capabilities:

a. Real-time error detection in commented practice phase 2 problems.

b. Student feedback after graded practice phase 3 problems in the form of annotated replay and performance summaries.

c. Instructor feedback emphasizing overall progress.

d. Adaptive problem selection.

e. Maintenance of student records.

The performance measurement subsystem (PMS) was designed to detect behaviors which do not conform to those correct behaviors described by the behavioral objectives. Table 9 shows the enabling behaviors which are monitored, categorized by terminal objective. It can be seen that both verbal and motor behaviors are monitored, as are omissions of required behaviors. The

TABLE 9. SKILL CATEGORIES

CATEGORY 1, HANDOFF

Controller Action	Partial Credit	Total Possible Points
A. Monitor feeder controller ICS	10	
B. Monitor proper frequency as specified in the handoff	10	
C. Acknowledge handoff		
1) Acknowledgment given prior to radar contact	10	
2) Acknowledgment given within 10 seconds	10	
D. Report radar contact		
1) Radar contact reported prior to radio check	10	
2) 50 percent of target on display at report	15	
3) Report not later than 10 seconds after 50 percent target appearance	15	
4) Call sign correct	5	
5) Radio frequency correct	5	
E. ICS off, radio frequency selected		
1) Pattern controller does not relinquish frequency, "Give me..." request made within 15 seconds	5	
2) Pattern controller relinquishes frequency and "Give me..." not used		
3) When pattern relinquishes frequency, ICS is deselected	5	
		100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 2, RADIO CHECK

Controller Action	Partial Credit	Total Possible Points
A. Radio Contact		
1) Within 30 seconds of 50 percent target appearance	10	
2) Proper frequency selected	10	
3) Mike keyed	10	
4) Call sign used	10	
5) One of the following given:	10	
a) "How do you hear..."		
b) "Wheels..."		
c) "Turn...heading"		
d) "Turn..."		
6) Mike unkeyed within three seconds and left unkeyed five seconds	20	
B. Speech quality		
1) Pilot responds "Loud and clear," or	30	
2) If pilot responds "Weak...,"		
a) Student answers "how...now," unkeys within three seconds and leaves unkeyed five seconds	15	
b) Pilot can respond "Loud...," i.e., V.U. level normal	15	
		100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 3, TURN-TO-FINAL

Controller Action	Partial Credit Turn	Partial Credit Straight-In	Total Possible Points
A. Accuracy of turn vectors, if given. (Score is given a weight of .6, score for B weighted .4; for a straight-in approach, the entire 100 points is given on B 1 and 2)			
1) Turn in proper direction	40		
2) Call sign correct	20		
B. Quality of turn or initial control			
1) At six miles (three for short approach) target is within two target widths of cursor	10	30	
2) At five miles (two short approach) target intercepts azimuth cursor in target zone one or two	20	70	
3) More than one turn used to turn aircraft onto final	10		
			100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 4, APPROACHING GLIDEPATH

Controller Action	Partial Credit	Total Possible Points
A. Approaching glidepath		
1) Transmission given	10	
2) Call sign and "over" needed and used	5	
Call sign and "over" not needed and not used		
3) Transmission given when aircraft is within the correct range	5	
4) Transmission given only once during final approach	5	
B. Do not acknowledge		
1) Transmission given only once	10	
2) Correct call sign used	5	
3) The phrase is not followed by "over"	5	
4) Transmitted prior to "begin descent"	5	
C. Begin descent		
1) Transmission given	10	
2) Transmitted within 10-30 seconds after "approaching glidepath"	5	
3) Glidepath cursor intersects upper 1/3 of target when advisory given	10	
4) Transmitted only once during the approach	5	

(Range table for item A.3:)

Aircraft Speed	Acceptable Range (Miles)
90	0.25-0.75
120	0.33-1.00
140	0.38-1.16
160	0.44-1.33
200	0.55-1.67

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 4, APPROACHING GLIDEPATH (CONT)

Controller Action	Partial Credit	Total Possible Points
D. Wheel check		
1) Transmission given prior to "approaching glidepath" when pilot has not said "wheels down"	15	
Transmission not given after pilot has said "wheels down"		
2) Correct call sign and "over" used	5	100

CATEGORY 5, HEADING TRANSMISSIONS

Controller Action	Weighting Factor Applied to Percentage Error	Total Possible Points
A. While range greater than five miles; all turns evenly divisible by 5°	.1	
B. Turns must not be of 1° o	.1	
C. All heading vectors		
1) Direction of the turn and heading digits correspond such that the direction advised causes the smaller turn	.2	
2) A counter-corrective turn made within eight seconds when a turn of more than 120° is given	.05	
3) Target enters zone three from zone two, a heading correction given within 20 seconds. This check is initiated when target has been in zones one or two for 1/2 mile, or at five miles (two for short approach), whichever comes first.	.15	
The heading given in the "Heading..." message the same as previously assigned	.25	
"Heading..." not used more than five times in an approach	.15	100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 6, AZIMUTH POSITION AND TREND

Controller Action	Weighting Factor Applied to Percentage Error	Total Possible Points
A. Position calls		
1) Position call correct	.5	
2) "Well" followed by a corrective turn within three seconds, or "correcting"	.25	
B. Trend calls		
"Correcting" used only when target is closing with centerline	.25	
		100

CATEGORY 7, GLIDEPATH POSITION AND TREND

Controller Action	Weighting Factor Applied to Percentage Error	Total Possible Points
A. For all glidepath messages, "begin descent" has been given	.10	
B. Position calls		
1) Position correct	.15	
2) A position call made whenever target changes zones, unless superseded by a priority call	.15	
C. Trend Calls		
1) Trend correct	.15	
2) Trend issued if the target moves from one zone to another	.15	
3) Trends not issued successively except in well zone	.15	
4) Trends do not separate identical position messages except in well zone	.15	
		100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 8, RANGE CALLS

Controller Action	Weighting Factor Applied to Percentage Error	Total Possible Points
A. All range calls made once the first one is made or five miles is reached, whichever comes first, unless superseded	.6	
B. The call made within ±0.1 mile of the mark	.2	
C. Correct range used	.2	
		100

CATEGORY 9, DECISION HEIGHT

Controller Action	Partial Credit	Total Possible Points
A. Decision height call		
1) Call given	25	
2) Target not touching cursors and call was followed by highest priority correct position	25	
B. Range		
1) DH announced within .80 miles from touchdown*	20	
2) DH announced prior to .70 miles from touchdown*	25	
C. Call is made only once during the approach	5	
		100

*Safety error

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 10, CLEARANCE

Controller Action	Partial Credit	Total Possible Points
A. Clearance requested		
1) Initial clearance request made after 3.1 miles	10	
2) Initial clearance request made prior to or at 2.9 miles	30	
3) Clearance not received and second request posted between 2.1 and 1.9 miles, or,	10	
Clearance received and not requested again		
B. Issuance of clearance when received from tower		
1) Correct wind information given	10	
2) Wind issued after clearance is received from tower	10	
3) Clearance issued after received from tower*	5	
4) Clearance issued after wind advisory	5	
5) Clearance issued prior to one mile	20	
or		
C. Clearance problems leading to a waveoff		
1) If clearance is not received		
a) Reason and waveoff issued prior to 1.3 miles*	35	
b) Proper missed approach transmission used.	15	
or		

TABLE 9. SKILL CATEGORY (CONT)

CATEGORY 10, CLEARANCE (CONT)

Controller Action	Partial Credit	Total Possible Points
2) If waveoff is given or clearance is cancelled		
a) Reason and waveof sued within two seconds r receipt of cancellation*	35	
b) Proper missed approach transmission used	15	
		100

*Safety error

CATEGORY 11, OVER LANDING THRESHOLD

Controller Action	Partial Credit	Total Possible Points
A. Over landing threshold		
1) Transmission given	20	
2) Given within ± one second of the target contacting the landing threshold point	20	
B. Final course position		
1) Given within three seconds of "over landing threshold"	20	
2) Position correct (including "over" for "on" position)	20	
3) "Over" is used correctly	20	
		100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 12, HANDOFF AND ROLLOUT

Controller Action	Partial Credit	Total Possible Points
A. Rollout instructions on full-stop landing		
1) Rollout instructions given	40	
2) Instructions issued 20-40 seconds after "over"	20	
3) Radio frequency is released within 10 seconds after rollout instructions	20	
4) Pattern controller is notified	20	
or		
B. Handoff to the pattern controller made if aircraft is on low approach or touch-and-go, or executing a missed approach including lost communications		
1) Handoff is given	40	
2) Handoff is made within 30 seconds of:	10	

Condition	Reference Point
Waveoff	Issuance of waveoff
Low approach	Decision height
Touch-and-go	Landing threshold

Controller Action	Partial Credit	Total Possible Points
3) Call sign correct	5	
4) Button correct	5	
5) If missed approach, range must be given to nearest 1/2 mile, else not		
6) Monitor frequency and ICS until pattern transmits "CS radar"	10	
7) Release radio frequency	10	
8) Pattern ICS selected during handoff	10	
		100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 13, NO-GYRO

Controller Action	Partial Credit	Total Possible Points
A. Warn pilot	20	
"Heading XXX" given if 1/4 mile elapses after a turn and less than a 2° change in course is observed		
B. Prepare for no-gyro		
1) No-gyro approach announced	30	
2) No-gyro approach announced if course correction is not taken within 1/2 mile	10	
3) The announcement issued prior to 3/4 mile from the point at which warning was issued	10	
C. Make 1/2 standard rate turns		
1) Transmission given	10	
2) Issued after begin descent, and no-gyro announcement	10	
3) Transmitted only once	10	
		100

CATEGORY 14, NO-GYRO HEADING CORRECTIONS

Controller Action	Weighting Factor Applied to Percentage Error	Total Possible Points
A. Turn was in correct direction	.4	
B. "Stop turn" issued	.4	
C. If target enters zone three from zone two a heading correction given within 20 seconds	.2	100

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 15, EMERGENCY WAVEOFFS

Controller Action	Partial Credit	Total Possible Points
A. Radar contact lost		
1) If target moves off the display or the display fails, waveoff issued*	50	
2) Issued within five seconds*	25	
3) Proper R/T used for type of approach	25	
or		
B. Target not touching at decision height		
1) Target not touching when decision height message given and waveoff issued*	50	
2) Followed by "Too low" message if aircraft was too low, else by some "too..." message. (Correctness of message scored in PV09, A2)*	25	
3) Proper R/T used for type of approach	25	
		100

*Safety error

TABLE 9. SKILL CATEGORIES (CONT)

CATEGORY 16, LOW ALTITUDE ALERT

Controller Actions	Partial Credit	Total Possible Points
A. Low altitude alert		
1) Transmitted when target exceeds one target width per mile below glidepath	50	
2) Issued within five seconds	50	
		100

CATEGORY 17, TRANSMISSION BREAK

Controller Actions	Weighting Factor Applied to Percentage Error	Total Possible Points
A. Mike unkeyed after "over"	.8	
B. At least one break given subsequent to "do not acknowledge" and prior to one mile	.2	100

CATEGORY 18, TRANSMISSION RATE

Controller Actions	Total Possible Points
A. Transmission rate after "do not acknowledge" advisory: Not more than five seconds between advisories	
	100

TABLE 9. SKLL CATEGORIES (CONT)

CATEGORY 19, RADAR ALIGNMENT

Controller Action	Partial Credit	Total Possible Points
A. Alignment check preparation		
1) Azimuth: servo down until center-line reflector appears	10	
2) Elevation and range: servo left until touchdown reflector appears	10	
B. Select ALIGN if alignment of		
1) Azimuth	20	
2) Elevation	20	
or		
3) Range	20	
is needed; else not		
C. Reposition antennae		
1) Servo up until one-mile mark is bisected by glideslope	10	
2) Servo right until the one-mile mark is bisected by azimuth cursor	10	100

A, B, and C must be performed sequentially or no credit is given first capability referred to above is provided by using PMS strictly as an event detector. Capabilities b through e are provided by the invocation of a scoring module which transforms the record of events into the numerical scores shown in the table.

These performance data requirements together with four crucial operational constraints provided the functional specification for the PMS. These constraints were:

a. That the system be capable of detecting errors in a particular terminal objective and of ignoring other errors;

b. That error detection take place in real time in phase 2;

c. That the effect of improper verbalizations used during phase 3 be nullified through the use of the word "correction;"

d. That the system be capable of re-scoring the performance test after the correction of misrecognitions.

The second and third constraints prevent performance measurement from working quite the same way in both phase 2 and phase 3 because in the former, error detection must be immediate. In the latter, it cannot be immediate else "correction" could not be used. On the other hand, the complexity of the subsystem as reflected in Table 9, made it clear from the outset that the development of two separate subsystems for phase 2 and 3 would not only be difficult, but would prove impossible to maintain. Therefore, a scheme was devised for invoking PMS whereby the phase of instruction is transparent to it. This was accomplished as follows.

First, PMS was designed to use the performance data placed in its own buffer, and to be unaffected by the way this buffer is filled. Then, since the data required for PMS are the same as those needed for replay, each of the major subsystems was designed to be responsible for reporting the specified activities to a central replay/PMS activity file. The data required are shown in Table 10. This reportage is all funneled through one reentrant subroutine. In phase 2, this subroutine causes the PMS buffer to be filled directly, as events occur. In phase 3, it causes the data to be saved on disk so that they can be read in after the run for use by PMS.

In phase 2, the PMS executive is called directly by the problem executive, and simply calls the PMS function-specific processors. After a phase 3 run or after modification of the performance test data, the PMS executive first reads the activity file data into the PMS buffer, and then processes the data by calling the function-specific processors. The "correction" processing is enabled by phase-specific processing in the speech understanding module which does not release a recognition record to the activity file in phase 3 until the next recognition is encountered. If it is "correction," an indicator is set in the recognition record. Thus, PMS is able to satisfy its diverse set of requirements, and the phase-specific processing is actually confined to only three routines.

TABLE 10. DATA SAVED IN THE REPLAY/PMS ACTIVITY FILE

Type of Record	Contents
Header	Initial conditions of the problem
Replay synchronization	Time record was written Digitized speech record number
Speech understanding	Time Range Recognitions (first and second choices) Mike key information Model controller selections Aircraft position, trend and heading
Trainee panel changes	Time Range Trainee panel status
Servo changes	Time Servo position
Automated voice output	Time Range Phrases spoken
Special scoring records	Time Event indicators for 50% of target appearing, aircraft entering zone three from zone two, mile mark passed, gyro failure, etc.

Once this theory of operation was established for PMS, the requirements for real-time operation in phase 2 and for terminal-objective-specific scoring, led to the design of a highly modular structure as opposed to a large, general, table-processing scheme. The latter was considered to be too expensive both in terms of processing time and core requirements.

A scheduling procedure was devised to provide an effective means of detecting range- and time-related events. Thus, there are no routines which tie up processor resources while waiting for such events. Instead, they are called upon as needed by the range and time scheduling routines. The time scheduler was designed to operate on the basis of apparent time so that PMS could operate in real time during a phase 2 problem to detect some condition such as the issuance of a turn within 20 seconds, and also could operate faster than real time to score an entire phase 3 problem after the run is completed within a few seconds of elapsed time.

Speech Recognition and Speech Understanding. The speech preprocessor is the device that makes it possible in theory to automate the training of the GCA controller because it enables the computer to acquire information about verbal performance. The preprocessor samples, filters, transforms and digitizes the speech signal. The rest of the speech understanding task is accomplished by software. The lessons learned in the laboratory system were used to develop diverse techniques to make speech recognition by a naïve user a practical reality (Hicklin and Slemon, 1978).

The recognition algorithm is speaker-dependent, that is, it performs recognition by comparing speech input to previously stored patterns. Figure 12, drawn from the Student Guide (Hicklin et al., 1980b), illustrates this point and explains why samples of the utterances must be collected before recognition is attempted. The GCA-CTS training package was designed to elicit and collect these samples in a natural way. The actual algorithm used for reference pattern creation is similar to that employed by the manufacturer of the speech recognition hardware, with one important exception. One of the difficulties with the recognition of the GCA-CTS vocabulary is that it consists of both long and short phrases. The need to distinguish between long similar phrases led to the implementation of a longer reference pattern in the

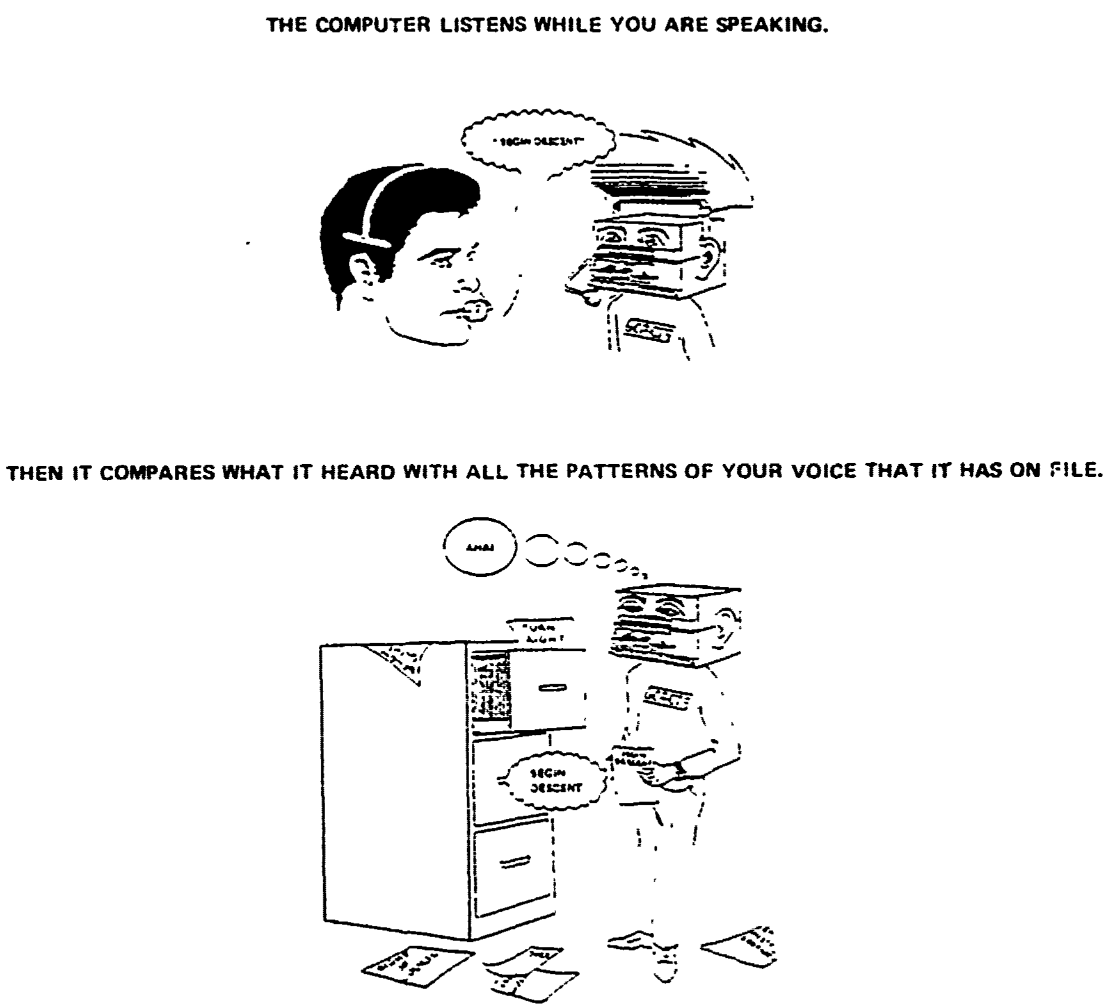

Figure 12. Speech Recognition in GCA-CTS

laboratory system. It was found, however, that recognition accuracy declined for the very short words such as the digits. To maintain high resolution for long phrases and to improve recognition for short phrases, two reference pattern lengths are used in GCA-CTS.

The phrase list recognized by the system is shown in Table 11. It was constructed to accommodate the PAR controller radio terminology while minimizing stylization requirements. Given the limitations of the state of the speech recognition art at the hardware procurement stage, some stylization is inevitable because recognition is only possible for words or phrases spoken in isolation. However, with the phrase list constructed in this way, the majority of the phrases are complete transmissions. This attempt to minimize stylization requirements produced a phrase list which places three very demanding requirements upon the speech recognition algorithm:

a. Many phrases are very similar, differing only in a small region;

b. The phrases differ vastly in length;

c. There is a relatively large number of phrases.

To satisfy the first requirement, a technique tested in the laboratory system was employed. It involves comparing the reference patterns for similar phrases to find the areas in which they differ significantly. This area in the speech input is then correlated with the same area in each of the reference patterns. This effectively causes the pattern recognition algorithm to weight the distinctive portion of the utterance more heavily than the similar portions. Analysis of recognition confusions between similar phrases revealed that in some cases, the variation in reference pattern density imposed a bias on the recognition algorith. Therefore, a scoring algorithm was devised to remove this source of bias.

The second requirement was satisfied by devising a scheme for collecting reference patterns of two sizes based upon the number of syllables in the utterance as described above. This further necessitated changes to the scoring algorithm to make the results of comparisons with the long and short patterns comparable.

To satisfy the third requirement, a vocabulary partitioning scheme was devised to restrict the branching factor. The recognition algorithm makes a first pass recognition attempt over the set of phrases the trainee is most likely to have said. If no recognition is found, a second pass is made over the other phrases which are valid in the particular phase of the approach. If necessary, a final pass is made over the rest of the vocabulary. This partitioning scheme is based upon the model controller selections and depends upon a special purpose speech level indicator built into the trainee panel. It works as follows. The model controller always has available a set of phrases which is correct at a given time. When the system detects that the trainee has started to speak, the set of legal phrases is sent to the recognition algorithm and forms the basis for the first pass vocabulary partitioning.

TABLE 11. GCA-CTS PHRASE LIST

PATTERN CONTROLLER DIALOG

POSITION FOUR ROGER
RADAR BUTTON ONE
RADAR BUTTON TWO
GIVE ME BUTTON ONE
GIVE ME BUTTON TWO
ON THE GO
MISSED APPROACH
1 MILE
1 AND 1/2 MILES
2 MILES
2 AND 1/2 MILES
3 MILES
3 AND 1/2 MILES
BUTTON ONE
BUTTON TWO
BUTTON ONE CLEAR
BUTTON TWO CLEAR

CALL SIGNS

ARMY EIGHT SEVEN SIX
MARINE SIX EIGHT SEVEN
NAVY THREE ONE ZERO
AIR FORCE THREE ZERO SEVEN
OVER

RADIO/WHEEL CHECK; APPROACHING GLIDEPATH SEQUENCE

THIS IS YOUR FINAL CONTROLLER HOW DO YOU HEAR ME?
HOW DO YOU HEAR ME NOW?
WHEELS SHOULD BE DOWN
APPROACHING GLIDEPATH
DO NOT ACKNOWLEDGE FURTHER TRANSMISSIONS
BEGIN DESCENT
OVER

TABLE 11. GCA-CTS PHRASE LIST (CONT)

RANGE CALLS

1 MILE FROM TOUCHDOWN
2 MILES FROM TOUCHDOWN
3 MILES FROM TOUCHDOWN
4 MILES FROM TOUCHDOWN
5 MILES FROM TOUCHDOWN
6 MILES FROM TOUCHDOWN
7 MILES FROM TOUCHDOWN
8 MILES FROM TOUCHDOWN

COURSE AND HEADING MESSAGES

ON COURSE
SLIGHTLY RIGHT OF COURSE
SLIGHTLY LEFT OF COURSE
RIGHT OF COURSE
LEFT OF COURSE
WELL RIGHT OF COURSE
WELL LEFT OF COURSE
CORRECTING
ON CENTERLINE
SLIGHTLY RIGHT OF CENTERLINE
SLIGHTLY LEFT OF CENTERLINE
RIGHT OF CENTERLINE
LEFT OF CENTERLINE
TURN RIGHT HEADING
TURN LEFT HEADING
HEADING
0
1
2
3
4
5
6
7
8
9

TABLE 11. GCA-CTS PHRASE LIST (CONT)

GLIDEPATH MESSAGES

ON GLIDEPATH
SLIGHTLY ABOVE GLIDEPATH
SLIGHTLY BELOW GLIDEPATH
ABOVE GLIDEPATH
BELOW GLIDEPATH
WELL ABOVE GLIDEPATH
WELL BELOW GLIDEPATH
COMING UP
COMING DOWN
GOING ABOVE GLIDEPATH
GOING BELOW GLIDEPATH
GOING FURTHER ABOVE GLIDEPATH
GOING FURTHER BELOW GLIDEPATH
AT DECISION HEIGHT

CLEARANCE

WIND
AT
0
1
2
3
4
5
6
7
8
9
CLEARED FOR LOW APPROACH
CLEARED FOR TOUCH AND GO
CLEARED TO LAND
TOWER CLEARANCE CANCELLED
TOWER CLEARANCE NOT RECEIVED

TABLE 11. GCA-CTS PHRASE LIST (CONT)

APPROACH TERMINATION

AT DECISION HEIGHT
OVER LANDING THRESHOLD
ON CENTERLINE
SLIGHTLY RIGHT OF CENTERLINE
SLIGHTLY LEFT OF CENTERLINE
RIGHT OF CENTERLINE
LEFT OF CENTERLINE
CONTACT TOWER AFTER LANDING

NO-GYRO PHRASEOLOGY

THIS WILL BE A NO-GYRO PAR APPROACH
MAKE HALF STANDARD RATE TURNS
TURN RIGHT
STOP TURN
TURN LEFT

UNUSUAL SITUATIONS AND WAVEOFFS

CORRECTION
TOWER CLEARANCE CANCELLED
TOWER CLEARANCE NOT RECEIVED
TOO LOW FOR SAFE APPROACH
TOO HIGH FOR SAFE APPROACH
TOO FAR LEFT FOR SAFE APPROACH
TOO FAR RIGHT FOR SAFE APPROACH
IF RUNWAY NOT IN SIGHT
IF RUNWAY NOT IN SIGHT EXECUTE MISSED APPROACH
EXECUTE MISSED APPROACH
CLIMB AND MAINTAIN ONE THOUSAND FIVE HUNDRED
TURN RIGHT HEADING
RADAR CONTACT LOST
CLIMB AND MAINTAIN THREE THOUSAND
TURN RIGHT
PROCEED DIRECT POINT BRAVO HOLD UNTIL ADVISED BY GCA
LOW ALTITUDE ALERT CHECK YOUR ALTITUDE IMMEDIATELY

OTHER PHRASEOLOGY

OVER
CORRECTION

For example, if the correct glidepath message is "below glidepath," the first pass compares the input to the phrases "on glidepath," "slightly below glidepath," "below glidepath," and "well below glidepath," as well as to a set of likely course and other messages. This effectively diminishes the probability of a system error in the form of a recognition confusion between the easily confused "above glidepath" and "below glidepath" transmissions, but allows the system to detect errors which the student is likely to make.

Speech recognition is only the first step in the process of making sense out of the verbal input. The second step is referred to as speech understanding because it combines the isolated phrase recognitions into complete transmissions (shown in Table 12) if necessary, and formats them in a way that makes them usable by the other subsystems. It also tests the recognitions for reasonableness and attempts to compensate for both trainee stylization errors and system recognition errors. As an example of the former, since the stylization requirements demand that the trainee pause between the digits in a heading command, there may be a tendency to pause unnecessarily between the digits in a call sign. If this mistake is made, the speech understanding subsystem will receive an unrecognized utterance followed by the three call sign digits. The speech understanding system interprets this as the aircraft call sign.

An example of the compensation for system recognition errors is the processing of turn commands. The speech understanding subsystem uses both recognition quality and environmental information to improve understanding. The speech recognition algorithm provides a measure of its confidence in the recognition, and includes its second choice if there is no significant difference in scores for two potential choices. This sometimes happens when a turn is given. When it does, speech understanding examines the environmental conditions to determine whether one of the recognitions is more correct than the other. If it is, the more correct phrase is assumed to have been used. Thus, the system gives the trainee the benefit of the doubt if there is any chance that a recognition error has occurred.

Aircraft/Pilot/Environmental Simulation (APE). APE was designed to provide a simulated environment in which the trainee can acquire and practice control skills. The requirements for the skill acquisition environment are quite different from those for the environment suited to practice. Considering the latter, during practice sessions, the simulation must provide a realistic environment to facilitate transfer of training. This does not mean, however, that such things as the performance characteristics of a variety of aircraft should be simulated — this is not observable to the PAR controller. What is observable to the controller is target response to verbal transmissions and wind effects. Therefore, the design focused upon making these aspects of the simulation realistic. This is accomplished in the following ways.

First, GCA-CTS provides the trainee with a high fidelity simulation of a PAR display. APE causes the simulated target to move across the display in a manner which closely approximates the motion of the actual PAR image of a real aircraft. The target's motion varies in response to trainee transmissions and

TABLE 12. GCA RADIO TRANSMISSIONS

Transmission

1) "Position 4 roger"

2) "C/S, radar button X" (C/S = call sign of aircraft)

3) "Give me button X"

4) "C/S, this is your final controller how do you hear me?"

5) "How do you hear me now?"

6) "C/S, turn right heading XXX, over"

7) "C/S, turn left heading XXX, over"

8) "Wheels should be down, over"

9) "On the go, C/S, button X."

10) "On glidepath"

11) "Slightly above glidepath"

12) "Slightly below glidepath"

13) "Above glidepath"

14) "Below glidepath"

15) "Well above glidepath"

16) "Well below glidepath"

17) "Coming up"

18) "Coming down"

19) "Going further above glidepath"

20) "Going further below glidepath"

21) "Going above glidepath"

22) "Going below glidepath"

23) "Tower clearance cancelled, execute missed approach, [climb and maintain one thousand five hundred, turn right heading 300]"

24) "Tower clearance not received, execute missed approach, [climb and maintain one thousand five hundred, turn right heading 300]"

25) "Heading XXX"

26) "C/S, approaching glidepath, over"

27) "Approaching glidepath"

28) "Begin descent"

29) "Missed approach, C/S, (map position), button X"

TABLE 12. GCA RADIO TRANSMISSIONS (CONT)

30) "Well right of course, turn left heading XXX"

31) "Well left of course, turn right heading XXX"

32) "Well right of course, correcting"

33) "Well left of course, correcting"

34) "Right of course"

35) "Left of course"

36) "Slightly right of course"

37) "Slightly left of course"

38) "On course"

39) "X mile(s) from touchdown"

40) "At decision height"

41) "At decision height, too X for safe approach, if runway not in sight, execute missed approach, [climb and maintain one thousand five hundred, turn right heading 300]"

X = 1) high
2) low
3) far right
4) far left

42) "Wind XXX at X, cleared X_1"

X_1 = 1) for low approach
2) for touch-and-go
3) to land

43) "C/S, do not acknowledge further transmissions"

44) "Over landing threshold, [X centerline], over"

X = 1) left of
2) slightly left of
3) on (may be omitted)
4) slightly right of
5) right of

45) "Over"

46) "Contact tower after landing, over"

47) "Button X, clear"

TABLE 12. GCA RADIO TRANSMISSIONS (CONT)

48) "Radar contact lost, [if runway not in sight] execute missed approach, climb and maintain three thousand, turn right, proceed direct point bravo hold until advised by GCA"

49) "Low altitude alert check your altitude immediately"

50) "C/S, this will be a no-gyro PAR approach, over"

51) "C/S, turn left, over"

52) "C/S, turn right, over"

53) "C/S, stop turn, over"

54) "This will be a no-gyro PAR approach"

55) "Make half standard rate turns"

56) "Turn left"

57) "Turn right"

58) "Stop turn"

other approach events in the same way as the target return from a real aircraft would.

The simulated pilot causes this target response by formulating a conception of the current correct rate-of-turn, rate-of-climb, and airspeed for the aircraft from the most recently received transmission or from the most recently encountered approach event, in accordance with a body of specific rules and procedures which dictate the proper behavior of pilots on final approach, and then by attempting (with a degree of success dependent on the pilot's skill level) to manipulate the controls of the aircraft to achieve and maintain the above correct rate-of-turn, rate-of-climb, and airspeed.

The simulated target's motion also varies in a manner similar to those motions of the PAR image of a real aircraft on GCA which appear attributable to the action of wind on the aircraft.

Secondly, the simulated pilot emits utterances (using computer-generated speech) in response to controller transmissions and other approach events. The simulated pilot's verbal behaviors duplicate the verbal behaviors of a real pilot conducting an actual approach.

Finally, the simulated PAR display includes two numbers, labeled "wind speed" and "wind direction," which vary with time in a manner which closely approximates real wind speed and wind direction time histories, and which correspond to the wind speed and wind direction required to produce the wind-induced effects described above.

In the broadest sense, APE acts within GCA-CTS as a "black box" to transform an input stream of controller advisories (which it receives aperiodically from the speech understanding subsystem) into an output stream of aircraft position vectors (which it sends each 0.5 second as inputs to the display-simulation routines), and pilot verbal reply specifications, in the form of phrase-identification numbers (which it sends aperiodically as inputs to the speech-generation routines).

This transformation performed by APE involves four separate but concurrent processes: the wind simulation, pilot thought process simulation, pilot verbal behavior simulation, and pilot motor behavior simulation.

The wind simulation models a wind of predetermined intensity, direction, and variability, blowing across the approach track. Wind is modeled as the sum of a steady component and a random component, modified by gusts. The random component is modeled as the combination of two uncorrelated processes acting along and across the steady wind vector. Each of these processes has an autocovariance as a function of time. The result of this autocorrelation is that successive samples of wind (taken each half second) are not independent, thereby preventing wild variations in wind velocity and direction. Gustiness is modeled as occasional increases and decreases in the wind, affecting the component along and across the steady wind direction equally. Gusts and "antigusts" (decreases in wind intensity) are assumed to occur equally frequently, with equal average durations. The steady wind speed is assumed to be equal to the geometric mean of the mean wind speed during gusts and antigusts along the steady wind direction.

The ongoing thought processes of the pilot are modeled by simulation of the pilot:

a. Detecting advisories embedded in the stream of incoming controller utterances;

b. Deciding (on the basis of a prolonged absence of incoming advisories) that he has lost radio contact with the controller and should execute a self-initiated waveoff;

c. Comprehending the content of a newly-detected advisory;

d. Monitoring the audibility of controller advisories and deciding whether he should verbally notify the controller that his transmissions are "weak but clear" as opposed to "loud and clear";

e. Deciding for any of a number of reasons, to execute a missed approach in spite of the fact that the controller has not issued an advisory to that effect;

f. Deciding what verbal reply, if any, to render in response to a newly-received advisory or a newly-encountered approach event of some other type, and deciding to delay rendering that reply until such time as it may be solicited by the controller uttering the word "over";

g. Reconceiving, each time the subroutine is called, a new current correct rate-of-turn, rate-of-climb, and airspeed for his aircraft, based on all information currently available to him (primarily, the most recently received advisory) and in accordance with all specified rules of pilot GCA behavior;

h. And thereafter, attempting to achieve and maintain the above correct rate-of-turn, rate-of-climb, and airspeed.

The pilot verbal behavior simulation was designed to transmit to the speech-generation routines a request to generate whichever verbal reply the pilot currently thinks is appropriate, if any.

The pilot motor behavior simulation models the pilot's actual (as opposed to attempted or correct) motor behavior and the dynamic response of the aircraft to the pilot's motor behavior. It outputs the current aircraft position vector to the display subsystem. The model works in the following way. First, the current true value of rate-of-turn, rate-of-climb, and airspeed are computed by applying to the current correct values of those variables certain error-inducing processes which embody, in a single, integrated (and indecomposable) model, both the pilot's skill level in achieving and/or maintaining any specific instrument picture he may desire, and the sensitivity of the dynamic response of the aircraft types being simulated to the pilot's motor behavior. Next, the aircraft's actual rate-of-turn is integrated with respect to the time required to determine the current aircraft heading with respect to the frame of reference of the simulation coordinate axes. Wind velocity, aircraft heading, and true airspeed are then used to determine the aircraft's current velocity with respect to the surrounding air mass. That velocity is then resolved into x-, y-, and z-axis components, to which are added, correspondingly, the x- and z-axis components of the current wind velocity, yielding the aircraft's current velocity with respect to the simulation coordinate system frame of reference. This velocity vector is integrated with respect to time over a 0.5 second period to generate a displacement vector which, when added to the last-computed aircraft position vector, yields the current aircraft position vector.

The foregoing dicussion described the design which enables APE to provide a realistic environment for the practice of controller skills. The ideal skill acquisition environment, on the other hand, may not be realistic at all. The GCA-CTS had two problems to solve in teaching controller skills, especially in the area of glidepath control. The first was to teach the trainee to use the complex radio terminology, and the second was to collect representative voice samples. The latter is extremely difficult because the phrases are very similar to begin with (see Table 11) and a plosive dominates the area of dissimilarity. Any effort therefore to emphasize "above" or "below" tends to make the plosive obliterate the critical dissimilarities in the input feature patterns. In order to solve these problems, recognizing that skill acquisition may be enhanced in an artificial instead of realistic environment, a restricted flight mode was designed. In this mode, the simulated pilot does not respond to controller transmissions, but instead flies the aircraft on a sinusoidal-like path through the contiguous position zones specified by the courseware designer. This offers two advantages. First, it allows the trainee to rehearse the complex phraseology in an orderly sequence in response to a realistic stimulus; secondly, it allows voice samples to be collected for only those positions on one side of the glidepath. This little trick enables the courseware designer to force the trainee to practice getting good speech recognition under circumstances which preclude the above/below recognition confusion. The more confidence the trainee develops in the system, the more natural his verbalizations are and consequently speech recognition difficulties are minimized.

Display Transformations. The provision of a realistic radar display was regarded as of critical importance. Some considerable delays and difficulties were encountered in the process of deriving display equations for the simulated radar display. The overall problem, simply stated, was to derive a set of equations for transforming real space coordinates to display screen coordinates. Two significant considerations in this problem influenced the direction of inquiry: the transformation equations should be simple, both for ease of coding and speed of execution, and the transformation should be "realistic" in the sense that it is representative of what appears on the operational gear. The latter consideration is particularly important as a training issue. In particular, the horizontal and vertical display scale factors are critical: they control magnification of error in the flight path angle, and consequently, imprecision in this area could adversely affect transfer of training.

The requirements of the display transformation as they were initially described were:

a. The range is logarithmic.

b. The target on both the azimuth and elevation displays is at the same horizontal coordinate.

c. The glideslope is a straight line in both the azimuth and elevation displays.

d. A pilot flying on a straight line toward the radar traces a straight line on both displays.

e. Horizontal flight near the glideslope at a range of nine miles appears on the display as an upward trace at an angle between 30 and 45 degrees.

f. Horizontal flight near the radar appears as an upward trace on the display.

g. Elevation angles on the display are multiplied by about 8, so that a display range of 48° to -8° corresponds to the real radar's scan limits of 6° to -1°. Similarly, for azimuth, angles are multiplied by about 2.7, so that a display range of 40.5° to -13.5° corresponds to a real radar's scan limits of 15° -5°.

Specific problems were encountered in the course of deriving the display tranformation equations based upon the requirements a through g, namely:

a. The logarithmic form (a) was inconsistent with the requirement that the glideslope be a straight line on both displays (c).

b. Using the best available data and a general logarithmic function to determine horizontal scaling, a very poor fit to the range scale was obtained.

These problems required the reevaluation of the set of requirements to choose which were real and which were inconsistent with reality. Thus, requirement (a) was discarded because it was not possible to use the logarithmic form and reproduce the actual display geometry. Requirement (d) also had to be discarded to achieve consistency in the azimuth display simulation.

Allocation of Processing. The allocation of processing to the two computers was a significant design effort, and although it was not addressed specifically in the <u>Design Report</u> (Barber et al., 1978), the results were reflected in that document in the design of the data structures and the Interprocessor Bus (IPB) communication scheme. The foregoing discussions reflect only selected highlights from the fairly complex design, but should suggest that the GCA-CTS has a good many real-time tasks to perform during a problem. The practical aspects of getting this processing done in two minicomputers with one shared moving head disk should not be minimized, hence, the topic is addressed briefly here. The design goals in the allocation of processing were:

a. To keep distinct functional elements self-contained within a processor;

b. To keep I/O processing requirements within the limits of processor capability;

c. To minimize interprocessor communication;

d. To compress disk access requirements to those which could be accommodated by the hardware.

The resulting allocation is shown in Table 13. Functionally, it can be described as a division into a training system controller and a speech and display processor. From the standpoint of number of functions performed, this allocation is obviously unbalanced, and resulted in a complex save file structure for the training system controller. However, it was necessary to divide the functions in this way in order to enable the system to handle the real-time processing burden. For example, in CPU 2, the display processor steals up to 25 percent of each processor cycle, while the speech recognition system must process interrupts from the preprocessor every 2.2 msec. In addition, the speech recognition system must exercise a time-consuming pattern matching algorithm to search the large phrase list for a match to an utterance while inputs from a new utterance are being collected.

TABLE 13. ALLOCATION OF PROCESSING

Training System Controller (CPU 1 - 64K)	Speech and Display Processor (CPU 2 - 96K)
Training control executive	Foreground:
Phase executives	Display processing
Replay	
Performance measurement	Background:
Speech understanding	Speech recognition
Aircraft/pilot/environment	Keyboard processing
Radar	IPB I/O
Model controller	
Speech digitizer I/O	
Trainee panel I/O	
Keyboard processing	
IPB I/O	

In CPU 1, the speech digitizer inputs data at a rate of 16,000 bits or 1,000 words per second. Although it is a data channel device and so interrupt processing is minimized, these data must be written to the disk at the same time that extensive overlay activity is occurring.

To enable these requirements to be met, use was made of all the features of Data General RDOS. In CPU 2 the functionally distinct display and speech recognition processes were allocated to different grounds. To accomplish real-time speech recognition, the reference patterns had to be core resident, but their large size required that they be stored outside the 32K user address space. They are mapped in as needed using the window mapping facility. Furthermore, the virtual overlay feature which maps code residing in extended memory into user address space is employed to reduce the disk I/O requirements to a minimum during a problem.

In CPU 1 the extra memory above 32K was used entirely for virtual overlays, and no foreground program exists. The time-critical routines are allocated to these areas and therefore no disk I/O is required before they can be called. Less time-critical modules are allocated to disk overlay areas.

Also in CPU 1, a unique use was made of the window mapping feature. Ordinarily, it is used to map a physical page containing data into user address space so that those data can be processed. In GCA-CTS, it is used to map the speech digitizer data area out so the user address space can be used for code. The two speech digitizer buffers require 2056 words of memory which is a significant portion of user address space. Therefore, this area is mapped into user address space only at system startup time. After the data channel map has been set up to include the buffers, the physical pages are mapped out and the logical pages are used for disk overlays. The extended read and write block Fortran library routines are used for transferring data between extended memory and the disk.

Facilities Report. A Facilities Report was prepared as an appendix to the Design Report (Barber et al., 1978). After inspecting the facilities at NATTC, site preparation requirements were identified and described.

DESIGN REVIEW/ILS CONFERENCE. The completed design was reviewed at Logicon by personnel from NAVTRAEQUIPCEN, NAVAIR and NATTC. Minor changes were suggested at this meeting and incorporated. The topic of integrated logistics support was also discussed. After reviewing the options carefully and presenting them, Logicon suggested a maintenance philosophy which relied primarily upon vendor-supplied maintenance agreements rather than upon training of Government personnel. The need for a daily operational readiness test for the system hardware was brought forward, and Logicon agreed to provide this feature.

DEMONSTRATION TEST PLAN PREPARATION. A test plan was developed midway through the implementation phase to guide acceptance testing. The plan was devised to demonstrate the various system components individually and also to demonstrate the integration of these components within the phases of instruction. The document was prepared in a workbook format.

TRAINER DEMONSTRATION REPORT PREPARATION. The originally proposed schedule for the completion of the software could not be met due to the increased scope of the project. Thus, only part of the demonstration was completed as planned in May, 1979. At this time, phase 1 was demonstrated. The remainder of the demonstrations were completed in August, 1979. The Demonstration Test Report Logicon, 1979b) describes these tests, the discrepancies noted and their resolution.

After the completion of the demonstration tests, Logicon hosted two students from NATTC and a group of about 20 others were invited to watch the students use the system. Many suggestions were made at this time and were incorporated to the extent possible prior to shipment.

TRAINING SYSTEM PACKAGE DEVELOPMENT. Implementation of the system involved integrating the Government-furnished equipment and vendor-supplied hardware, building and installing the special purpose hardware, courseware development and integration, software implementation and integration, shipping and installation, and document publication.

Integration of Commercially-Supplied Hardware. Integration of hardware into a functional system can prove to be a time-consuming task and snags seem to be inevitable.

System Hardware. There were several problems related to the system hardware which impacted the development effort. These included delivery delays which ranged from a few weeks to seven months. There were various types of hardware difficulties ranging from improper configuration of components by the manufacturers to malfunctions arising during system use. Although most of these malfunctions arose while the system components were under warranty or covered by service agreements, still programmer time was required to help isolate or troubleshoot the problems, and the system down-time affected other programmers. In one situation, a minor malfunction was discovered which was traced to a design flaw and which was not corrected for over a year. Fortunately this problem did not seriously affect system use.

Furniture. The desks which were originally proposed for GCA-CTS were chosen as a cost-effective alternative to a look-alike PAR console. The look-alike was not deemed absolutely necessary to ensure transfer of training, and its development expense was not considered to be just ied in the experimental prototype. Simple desks were therefore acquired. However, the change in focus brought about by the instructorless learning philosophy brought with it changes in the function of system components, and the trainee desk arrangement has proven to be less than satisfactory. In particular, the CRT and keyboard were originally intended to be used infrequently and so were located out of the way. In the instructorless system, the CRT is used extensively for the presentation of textual information explaining presentations on the graphics display. Had this been envisioned at proposal time, a rack-mounted system for the graphics and CRT would have been proposed in which both could be viewed with a minimum of head twisting.

Installation of Special Purpose Hardware. No major problems were encountered in the manufacture and installation of the trainee panel, instructor panel, junction panel, speech digitizer and interface. All of these components have operated without failure since installed.

Courseware Development and Integration. The development of courseware for an instructorless training system was a significant effort not envisioned at proposal time. It was approached in the following way. The syllabus was first refined and finalized. The decision was made to provide a Student Guide (Hicklin et al., 1980b) as the primary information source, to extract the crucial elements of each task from it and to present these in a Computer Aided Instruction (CAI) format while collecting voice reference patterns. Then practice problems would be provided in which the trainee could practice and integrate the new skills with the old.

The Student Guide (Hicklin et al., 1980b) preparation involved the collection and arrangement of information, then verification and integration of source material. Finally, the document was revised extensively to improve its instructional value and copious illustrations were prepared. The resulting document looks imposingly thick, but the actual amount of reading (which is interspersed with system time) is relatively small. Four different colors of pages are used to help the trainee distinguish the content quickly. The "must know" information is assembled on color-coded and edge-indexed summary sheets at the end of every lesson.

The efforts of a subcontractor were secured to develop the CAI materials for three levels of instruction based upon the preliminary version of the Guide. Although their work was extremely helpful, it ultimately fell to the project programmers to revise these materials, to format them as required by the phase 1 executive, and to debug the courseware. There were several reasons for this. First, contract resources for this activity were very slim because, strictly speaking, the ability to provide CAI was not at issue in the experimental prototype. Secondly, in the exceedingly short time frame and in the absence of a completed system, it proved to be difficult to convey the power, flexibility and limitations of this unique multi-media system to instructional designers who were familiar with more traditional systems. Finally, contract resources precluded the development of a user-oriented instructional design language with error-checking features; thus the translation of the material was laborious and debugging was as significant an effort as the debugging of a computer program.

To accommodate the requirements within these very real time and money constraints, it was decided to make levels 2 and 3 "showcase" examples of the sort of training which could be provided, and to implement minimal CAI in the other levels. This decision ensured that the degree of CAI believed to be necessary during system familiarization could be provided, enabled the courseware to be completed in a timely way, and allowed for the possibility of evaluating the desirability and effectiveness of providing elaborate CAI experiences as compared to a minimum of CAI and an emphasis upon learning by doing in phase 3.

Software Implementation and Integration. The software development effort was increased in scope significantly by the shift to an instructorless learning concept, and resulted in the need for a three month time extension in the development effort. The philosophy behind the development did not change, however. While the initial coding was taking place, the system software was installed and modified as needed. This involved implementing a system device driver for the Votrax, configuring the operating system to include it, and converting the Fortran IV unmapped graphics package for operation in a Fortran 5, mapped environment. Integration then commenced almost immediately. Implementation was planned to provide software for integration in a top-down manner. Simulators were devised for large system elements where needed. Thus, for example, an APE simulator was developed to provide the inputs needed to check out the display processor, and a speech recognition simulator was implemented to accept "recognitions" from the keyboard to enable the checkout of the speech understanding system, and APE responses. Each programmer was responsible for coding and unit testing his or her routines, for releasing them along with any dummy routines required, and for supporting the integration of the modules into the GCA-CTS structure. Some of the highlights and frustrations in this activity are described below.

Pilot Simulation. Detailed investigation of pilot actions as they influence the training of approach controllers revealed the desirability of changing the initial design of the APE simulation. Fortunately, the indicated change was a simplification of the previous design, so that time spent in reformulating the design was offset by a shorter implementation cycle.

The change in design resulted from a more intensive investigation of three kinds of information than was possible during the initial design phase. The information received was:

a. Recommended pilot responses to controller actions;

b. Opinions of experienced pilots about actual pilot responses; and,

c. The influence of simulated pilot response on the controller trainee.

It was initially assumed that to simulate the behavior of an excellent pilot it would be necessary to carefully integrate the information provided to him by the controller in several messages in both the pitch and lateral control channels. This assumption was based on the interpretation of the controlled approach as a process of building a cumulative concept in the pilot's mind of a line in space to which he attempts to maneuver the airplane. There was a further underlying assumption that the controller advisories, taken individually without integration by the pilot, would not result in a smooth, well-controlled approach.

Further analysis revealed that both recommended and actual pilot response to controller information are inconsistent with the concept that pilot action is mediated by an accumulated, integrated and maintained concept of the line in space which he is to follow. His response is always most strongly (if not exclusively) influenced by the most recent controller-generated information, and the integration of this information plays an insignificant role. Contrary to the earlier assumption, a well-controlled approach can always be accomplished (in the absence of extreme weather, equipment malfunctions, etc.) without long-term information integration by the pilot.

The behavior just described is characteristic of a skilled pilot responding to an experienced controller. A skilled pilot will respond in a different manner to an inexperienced controller who makes errors, uses non-standard procedures or otherwise degrades the pilot's confidence in the controller's ability. In this case, the pilot might be expected to rely much more on his accumulated concept of the desired path in space and achieve a successful approach "in spite of" the controller. However, for training purposes it was considered to be most desirable to subject even the relatively inexperienced trainee controller to a simulated pilot who is heavily reliant upon the trainee's advisories, as there is a danger that a more autonomous simulated pilot might engender an attitude that a few isolated errors in his actions may be smoothed over by a competent pilot.

This approach is consistent with the philosophy followed in syllabus design, which presents the trainee with a graded sequence of tasks of increasing, but manageable, difficulty. If it had not been possible or desirable to develop such a graded sequence, it might have been appropriate to expose the trainee to a relatively self-sufficient simulated pilot initially, and reduce the pilot's autonomy — increasing his dependence upon the controller — in a sequence of graded steps. Previous experience in teaching this task indicates, however, that this approach to training is not necessary.

The published APE design included serial procedures for simulating separately the long term integration of information in the pilot's mind, and the pilot's controlling the aircraft to conform with his accumulated concept of where the glideslope was located. The actual simulation was simplified and integrated by eliminating the separation between pilot cognitive and control functions, and the interconnecting concept of the line in space.

The new pilot model is directly influenced by the contents of the most recently received controller information. In the pitch plane the inference is a desired rate of descent, and in the lateral plane it is a desired heading or rate of turn. The aircraft/pilot dynamics model was not substantially changed, nor were pilot skill factors.

Display Presentation. Sweep is an aspect of the display which was not considered to have an impact on transfer of training, but which does contribute to face validity. Although not contractually required, a realistic sweep simulation was devised, using the graphics library routines. The processing requirements turned out to be too great to be accommodated in real time. While it is possible that special purpose graphics routines could be devised to support sweep processing, this was not done, and no sweep display is provided. However, target position is updated at the sweep rate.

The debugging of the display routines presented a special challenge because they were designed to operate in the foreground, and no foreground CRT was available. Thus, for example, Fortran error messages would be lost when the system was operating in the foreground. To solve this problem and, more importantly, to save the data base generated in the dynamic environment to test the effect of error correction measures, a data acquisition scheme was devised. The background routine which conveys information to the foreground has optionally compiled code to write these same data to a file. The foreground executive also has optionally compiled code which causes it to ask which ground it is operating in at system initiation time. If operating in the background, it automatically retrieves its input from the data file instead of from the interground communications area. This made all errors reproducible, and enabled the software to be debugged using the symbolic debugger. Furthermore, the training system controller could be used for another purpose during the debugging activity.

Controller Models. The controller models (pattern and final) had a diverse set of requirements to satisfy, and some integration difficulties arose as a result. As an example, the final controller model must provide a set of correct phrases to the speech recognition algorithm. It also must select the most correct phrase for utterance during demonstrations. Obviously, a priority scheme was required. Further, since there are phrases which are theoretically correct within a certain range window, but which are ideally spoken at the midpoint of the window, the algorithm cannot operate on a strictly priority basis. Narrowing the window doesn't help because a long phrase could be chosen just before the window opened and could continue until after the window closed. So the algorithm had to include the capability to look ahead and decide to be quiet until it is time to say a phrase.

The slowness of the synthesized speech proved to be an insoluble problem. The controller model will occasionally have to skip a glidepath transmission simply because the aircraft will transit an entire zone while a previous message is being output. Course control is also adversely affected, especially when the pilot is acknowledging the transmissions because it is sometimes impossible to give a turn quickly enough. An obvious solution would be to use the existing option of having the model controller speak with a human voice via the speech digitizer. This would have been implemented had there been sufficient disk space for the requisite speech data files on the fixed disk.

System Software. The computer manufacturer's system software proved to be dependable and to provide all the promised features. Some problems arose, however. For example, the large number of entries in the training system controller's load-on-call table caused an overflow at assembly time, and a run time patch to GCA-CTS had to be implemented to work around the problem. Also, an undocumented feature of the load-on-call mechanism caused the range and time executives to lapse into an infinite loop on return from load-on-call routines. This necessitated a minor redesign of the executives.

A Big System in a Small Machine. The most significant integration hurdle was, quite simply, memory. Given the word size of the computer, only 32K words of memory are directly addressable at any one time. Fortunately, the operating system does not require to be resident in user space, but Fortran library routines must be. Stack space for each task and common data storage further reduce the memory available for user code. The problems related to memory size were manifested in several ways. First, "core crunching" commenced almost immediately. It wasn't long before the core limitations forced the exclusion of the debugger and symbol table. This loss of the ability to test changes with on-line patches was quite significant, especially considering the fact that the creation of each new executable program file required 15 minutes. Finally, to fit the large number of parallel functions into memory necessitated a complex overlay structure in the training system controller which was a challenge to integrate.

Shipping and Installation. Following the successful completion of the demonstration tests, the system was packed and shipped to NATTC, where it arrived in late August, 1979. The hardware installation proceeded with only minor difficulties. These difficulties arose because of hardware failures undoubtedly precipitated by the move. In particular, a 32K memory board in CPU 2 failed and had to be replaced, and one monitor developed the habit of causing keyboard failures. These problems were quickly resolved. Throughout the installation effort, NATTC personnel and the Education and Training Support Group were extremely helpful, provided invaluable assistance, and made a tremendous effort to learn as much as possible about the system.

Hardware installation was followed by software and courseware installation. The instructors immediately noted discrepancies between their existing training and the behavioral objectives identified for the GCA-CTS in early 1978. To the extent possible, changes were made to the GCA-CTS to reflect the current syllabus while the installation team was in Memphis. Also, there were minor problems which arose as the system was exercised by new users. Again, these problems were corrected by the installation team.

Informal training sessions were conducted for the users of the GCA-CTS, and the instructors were given hands-on experience while Logicon was present to answer the questions which arose. The training effectiveness testing procedures and system reliability record maintenance procedures were also described.

Document Publication. The documentation which was included in the training system package included vendor-supplied documentation, the GCA-CTS System Documentation, the Student Guide and Instructor Guide. The bulk of the vendor-supplied documentation was related to the computational system hardware and software. Manuals for all the peripherals were also secured with the exception of a voice input preprocessor manual. The documentation is assembled in binders and housed in a media cabinet.

The System Documentation (Barber et al., 1980) is a 660-page document which describes the GCA-CTS hardware and software. Narrative descriptions and block diagrams of each subsystem are provided, and module descriptions are given for each routine. All data structures are defined and a separately published cross-reference listing of all COMMON variables is also provided.

The Student Guide (Hicklin et al., 1980b) preparation was discussed previously. The Instructor Guide (Hicklin et al., 1980a) is a complementary 150-page document which explains the purpose and use of the system.

TRAINING EFFECTIVENESS TEST PLAN PREPARATION. A Training Effectiveness Test Plan was developed as specified in the contract. However, because of the fact that the Government had already contracted with others to perform many aspects of this study, and because the contractually agreed upon time frame could not be accommodated, the Government requested that the plan be revised. It was also requested that the topic of reliability be addressed. These modifications were made.

INTERIM SUPPORT AND ON-CALL SERVICE. Several trips to the school have been made to provide support. In November, some system problems arose which seriously affected reliability. Because they were due to a combination of problems, it took some time to straighten them out. When the system faults began, hardware trouble was suspected and a power monitor was installed. In early November, we went to Memphis to study the problem. Coincidentally, this trip was during a period of heavy thunderstorm activity. A number of power fluctuations which were severe enough to cause system crashes were recorded. Unfortunately, no log of system faults had been kept during the time that the power monitor was installed, so there was no way to correlate system faults with the power fluctuations. However, during the week at NATTC system faults were observed which occurred in synchrony with 15G19 and radar failures, therefore power problems were strongly suspected to be the cause of the GCA-CTS system failures. Before recommending a solution, however, the exact correlation between system crashes and power fluctuations had to be determined; therefore, it was requested that the power monitor be reinstalled. An examination of a two-week recording from this monitor revealed line fluctuations, but a study of the Data General documentation revealed that the equipment tolerances are such that the fluctuations should not have caused the system to fail.

During this time, system crashes were also observed when the GCA-CTS program was not running (some computer games were also delivered with the GCA-CTS). In time, the following hardware problems were discovered:

a. An important component of the disk drive which senses the temperature differential between the two disks had inadvertently been left disconnected;

b. The synchros which control the disk heads were badly out of alignment and, in fact, at the very edge of the published tolerances;

c. There was an intermittent problem that was ultimately traced to a CPU board in the training system controller.

In addition, an assembly language coding error which accounted for some of the problems was also found. When these hardware and software problems were fixed, system reliability improved tremendously.

Other support activities included the development of a short syllabus for use in a study to be conducted by Canyon Research Group, Inc. Modifications to the GCA-CTS were also required to support various requirements for their study. With the contracting officer's approval, these changes were made despite the interruption to Logicon's training effectiveness testing.

Finally, the users suggested a set of enhancements to the system to ma the GCA-CTS training course more nearly reflect the current training in the school. As many of these enhancements as could be provided within the allotted time were implemented and installed at the end of January, 1980. They included the reporting of safety errors and correction of procedures used when the tower clearance is either cancelled or not given.

SECTION III

RESULTS AND CONCLUSIONS

This section describes the results of the GCA-CTS project, including a discussion of the validity of the assumptions, an evaluation of the technology, and remarks about system effectiveness. No separate subsection is devoted to speculative observations. Instead, these observations are woven into the discussions which follow as appropriate.

ASSUMPTION VALIDITY

This development effort was not intended to perform traditional validation studies. Nonetheless the validity of the assumptions has been a matter of intense concern and informal observations have been made throughout the project and especially at NATTC. These are described in the following paragraphs.

TRAINING COURSE. Informal observations indicate that the trainees have been able to learn to use the system and have been able to master GCA control skills through the use of the system. Diverse feedback has been received about the courseware. Some think that all the levels should be as complete as levels 2 and 3. Some think that there is too much textual information in these levels. But the consensus seems to be that the integration of voice data collection and the use of demonstrations in these levels is helpful.

AUTOMATED INSTRUCTOR. The automated instructor is able to direct the student's progress such that, by the performance test, successful performance is being displayed. A complaint was lodged against the problem adaptation logic in the automated instructor which should be discussed. If the trainee is having difficulty as reflected in low scores, this logic makes specific aspects of the problem easier. One of the ways it does this (depending upon the skill which isn't being performed properly) is to choose the slowest aircraft. Neither the trainees nor the instructors liked this because the slowest aircraft takes a very long time to make an approach. Thus we were asked to disable this aspect of the adaptation logic. Really a better solution would be for the automated instructor to choose a medium speed aircraft rather than the slow one. Perhaps the slowest aircraft used in GCA-CTS should be used only rarely, for example when the trainee is learning to deal with crosswinds.

The remedial training selection feature of the automated instructor was insufficiently exercised to make a statement about the validity of the approach used. This logic enables the automated instructor to select a remedial exercise specific to the skill in which the trainee is having trouble, and at a level consonant with his present level of knowledge. Thus a wide variety of remedial training exercises could be provided. They could include voice data collection to update voice patterns in case the difficulty is due to changes in the trainee's speech patterns. Unfortunately, contract resources did not permit the development of a separate set of remedial exercises. However, remedial tasks from the syllabus were specified in the original courseware delivery. The reduction in training time to five half days (from an original eight full days) caused many courseware cutbacks however, and all remedial tasks were removed except those specified for level 6 just prior to the performance test.

PERFORMANCE MEASUREMENT. The performance measurement logic does provide the required data for trainee and instructor feedback and adaptive course control. In the enhancement phase this logic was augmented to distinguish between safety errors and less serious errors.

EFFECTIVE FEEDBACK. Efforts to provide effective feedback were successful, but there is room for improvement in this area and many suggestions have been offered. First, a replay with error reporting takes a long time and is not used when there are relatively few errors reported. Thus, many times a trainee will remain unsure of exactly why the system says a particular skill area "needs work." An intermediate level of feedback is needed similar to the task performance printout given to the instructor. It could be implemented with the present equipment suite by enabling the trainee to request a report of the errors made in a particular skill category.

Another difficulty arose because the performance measurement subsystem is extremely observant and detects many subtle errors. Every effort was made to describe these to the trainee in a meaningful and non-repetitious way. For example, associated with every possible error is a set of three verbal statements. The first two are similar but differently worded statements about the nature of the error. One of these is always chosen. A third statement amplifies the discussion and is either always or randomly appended to the first statement. Other information is also presented if it is appropriate such as the message understood and the correct transmission. In observing the system in use, it appears that in spite of all this, some of the error explanations do not convey the information to the trainee adequately. The file of these explanations is very easy to change, and many improvements have been made, but some of the explanations could still stand to be polished up.

Another feedback problem arose as a result of the fact that the system is an experimental prototype designed to evaluate certain specific aspects of the technology. One feature of an operational training system which it does not have is logic to detect that the trainee is trying to beat the system. The contract resources were judged better spent in other areas. Thus it was expressly stated in the Design Report (Barber et al., 1978) that the system was designed for "motivated and responsible students." It is therefore possible to beat the system and get feedback saying "perfect." For example, course position information is of an advisory nature only and is not required to be used. If the trainee does not give any course position messages, a feedback statement of "Perfect" will be given when a better statement would reflect the fact that no use was made of the terminology at all. Obviously in an operational training system this would be prevented. The logic to prevent it is simple, but of significant magnitude because it must be designed specifically for each skill category and must be implemented in such a way that the selection of remedial training is not affected.

The commented practice phase 2 problems were designed in an attempt to extinguish improper responses through providing immediate feedback. Because the laboratory system revealed that this feedback was disruptive to the train of thought, problems are restarted (up to two times) rather than continued after an error. This solution really is not ideal since the trainee may not get through the entire procedure in three tries. It has been suggested that a printout be provided which gives the last messages spoken and that the problem continue after an error. Another possibility might be to replay a previous portion of the problem up to the error, then let the trainee perform the procedure correctly and continue. Finally, since the experimental prototype is much smarter than the laboratory system was insofar as it is capable of freezing and reporting only those errors in a specific skill category, it is possible that simply continuing after the error would not be so difficult for the trainee as it proved to be in the laboratory system.

One further aspect of feedback should be mentioned. It may be that the system gives the trainee too much information about a problem in some respects. It was specifically requested that the call sign and approach type be included in the problem description which is written to the CRT. It was argued that after the problem begins, the trainee's attention would be directed away from the CRT, and yet the information would be available if it was absolutely necessary to refer to it. Casual observation suggests, however, that the trainees are not listening to the handoff message as they should be, and are relying entirely on this presentation. This is undesirable because they are failing to learn the essential skill of listening for and remembering this information and also are learning to turn their attention away from the radar display.

OTHER INSTRUCTIONAL ASSUMPTIONS. No transfer of training studies have been conducted to date, so it is not possible to evaluate the impact of stylization constraints or the use of a simulated radar. In addition, the need for a Student Guide has not been evaluated although this could be done by comparing the performance of students who used it with that of students who did not.

TECHNOLOGICAL ASSUMPTIONS. The technological assumptions were in general supported, though it is safe to say that the system resources are taxed almost to their limit to support GCA-CTS. For example, a sweep simulation was regarded as a feature which would add face validity and which could be added in later stages of training once the trainee had learned to detect the relevant stimuli in the radar display. Although it was not a required contractual item, a sweep simulation was written. It was found that the processing for a radar display with sweep and speech recognition could not be managed in real time in the same computer. The problem was not analyzed to determine a solution because sweep was a very low priority feature. Still it is known that to include it would require at least the development of some more efficient graphics library routines and perhaps would require different graphics hardware.

The system disk memory resources were barely sufficient. The courseware and executable files take up most of the room on the fixed disk and there is room for only one set of student files per removable cartridge disk, primarily because of the enormous amount of digitized speech data which must be saved. More serious is the fact that the disk I/O burden during a phase 3 problem is very heavy. Any significant modification to GCA-CTS would definitely require a faster and larger disk to support it.

An additional 32K of memory was added to the speech recognition and graphics display processor during the course of the project. With the expanded (96K words) memory, it was possible to have the trainee speech data and all the speech recognition logic resident during a problem. All graphics routines are likewise resident in the foreground. This means that virtually no disk overlay activity goes on in this processor during a problem, and the shared disk is freed for almost exclusive use by the training system controller.

The vendor-supplied software proved to be reliable and in general provided all specified capabilities. The Fortran 5 language, though not structured, does have very powerful multitasking features which made it the language of choice at the commencement of the project. The compiler generates optimized code and takes advantage of all the features of the Eclipse hardware. Careful coding practices in conformity with the strengths of the language resulted in code that is easy to debug and maintain.

MANAGEMENT ASSUMPTIONS. Some diverse points with regard to the management decisions made during the project should be mentioned. First, although the documents produced did not conform to a predefined format, each one proved to be extremely useful. None were written simply to satisfy a contractual requirement. Instead, each was written to document the culmination of a well defined phase of the effort and to serve as the basis for the next phase. It was precisely because of this attitude that the documents proved to be useful in terms of both serving as baseline documents and accurately conveying progress to the Government. Further, this project demonstrated that with proper planning, it is possible to deliver software descriptions and user guides with the system instead of months later. It also demonstrated that useful, accurate documentation is not inexpensive in terms of publication costs, but well worth it in terms of the communication it facilitates.

A potentially risky management decision was made, namely to gain assistance from a subcontractor in the development of the CAI materials. Their work served as an important baseline for the development of the textual materials presented in levels 2 and 3. However, within the small budget allotted for this effort it was not possible for them to develop the actual courseware specifications. A very interesting feature of the GCA-CTS effort emerged during this endeavor which should be brought forward. We who have been working in the area of applying the automated speech technologies, both in general and specifically to the GCA-CTS task have acquired a depth of understanding which was simply not possible to convey fully even to highly qualified instructional technologists in the brief course of a couple of months. The attempt placed a more significant burden upon the in-house subject matter experts than was expected. The instructional design strategy ultimately was our own, although the subcontractor provided invaluable assistance. The strategy sprang from a really clear understanding of both the GCA controller task and the strengths and limitations of the technology. Although it may seem very strange to have programmers developing courseware, perhaps with a system which is employing state-of-the-art technology it is actually necessary for them to generate the prototypical materials because it is only they who really understand all the resources which have been built into the system in an effort to make the technology work. Naturally this approach would never work unless the persons involved had multidisciplinary skills.

Finally, many management decisions with regard to staffing the project required revision. With the expansion to an instructorless training concept came a need for a correlative staff expansion, but it was felt that this need would decline during integration. Not so. The more complex system required a significantly longer integration effort than was planned. Likewise, the on call support originally specified, proposed, and budgeted was simply not sufficient for a system of GCA-CTS' evolved complexity.

TECHNOLOGY EVALUATION

The project was designed to evaluate the automated speech technologies in an operational environment. The speech recognition system is evaluated in the next subsection. The other two devices, the Votrax Speech Synthesizer and the Logicon-designed speech record/playback device or speech digitizer, both proved to be reliable hardware wise. The Votrax seemed to be intelligible to all who heard it, and the digitizer was able to provide an intelligible replay of the trainee speech.

Two problems arose with respect to the use of the speech synthesizer and should be mentioned. First of all, a fairly common misconception among instructors not associated with the project was that the system was designed to teach the trainee to talk "like Egor," that is, like the Votrax. This notion could not possibly contribute to acceptance of the system by the instructor cadre.

The second difficulty arose in that at least one student was observed to mimic the Votrax during voice data collection. The instructional materials elicit speech patterns in a variety of ways precisely because the potential for this problem has been noted. In addition, the students are cautioned to imagine the control situation and to speak naturally. Apparently not all trainees do this. Any unnatural vocalization during voice data collection is virtually sure to preclude good recognition.

EFFECTIVENESS EVALUATION

A training effectiveness test plan was prepared during the course of the project, but the need for such an evaluation was diminished after the Government awarded a separate contract for the evaluation of the GCA-CTS. Nonetheless, an attempt was made to conduct the evaluation in terms of recognition system reliabilty and overall system reliability. The evaluation period began immediately after delivery during the time before all the hardware and software problems were worked out, and when the instructors were just becoming familiar with the system. Although the required ten students started the training course, several were transferred to the regular course before they could complete training. The amount of time spent using the system for each of these trainees is shown in Table 14. Four of the trainees actually finished the course, but some records for these students were not available to Logicon. The training effectiveness testing period was not extended to gain more data because the Government wanted to free the system for use in studies by another contractor.

TABLE 14. STUDENTS USING THE GCA-CTS SYLLABUS DURING TRAINING EFFECTIVENESS PERIOD

Student	Days Spent Using GCA-CTS	Performance Test Completed
BA	4	yes
BB	3	no
EL	3	no
KM	4	yes*
MN	5	yes**
NN	4	no
LR	1	no
?R	?	yes***
SS	1	no
BB	3	no

* performance test records not available
** no annotated performance test printout saved
*** no records saved except performance test printout

RECOGNITION SYSTEM RELIABILITY. The method defined to be used for assessing speech system reliability was to ask the instructors to observe a performance test replay with the performance test printout in hand, and to note any discrepancies between what the system reported having recognized and what the trainee actually said. (This same procedure is routinely used to enable the performance test results to be modified to correct any recognition errors.) The performance test is considered to be a reasonably demanding test for the speech recognition system for two reasons. First, the system is attempting recognition over virtually the entire phrase list (excluding nine phrases used only in the enrichment level).

Secondly, the final examination setting puts as much pressure on the trainee as any he or she is likely to encounter in the course of GCA-CTS training. Since stress is known to have an adverse effect on recognition accuracy, at least for some talkers, the performance test should provide a worst-case example of the trainee's recognition accuracy.

In order to attempt to make some statement about SUS reliability, all the records which were supplied to us have been studied. This includes the two sets of records retrieved during the actual testing period as well as four sets of records from trainees who completed the course after the other contractor's studies were complete. The sample size is still very small, but some interesting regularities are evident.

First of all, Table 15 presents the actual recognition errors observed for each of the trainees. The question marks in the "phrase(s) recognized" column indicate unrecognized phrases. Some of these unrecognized phrases can be attributed to stylization errors because the recognition system could not detect the required interphrase pause. The cause of the remainder cannot be precisely determined from the available data. There are unique consistencies within each particular trainee's performance. For example, glidepath messages were poorly recognized for trainee A. Trainee R apparently tended to forget to pause between the digits in heading commands. The only trainee for whom the system confused "turn right" and "turn left" was trainee B.

These data are summarized in Table 16. In this table, the term "phrase" refers to a recognition unit as defined in Table 11. It is clear from the table that the percentage of total phrases which is misrecognized is relatively small and that there is relatively little variability in misrecognition rate between speakers. The percentage of unrecognized phrases, on the other hand, varies by an order of magnitude between speakers. There are many potential reasons for this. Perhaps the extremely high rejection rate seen with trainees A and R is due in part to the fact that the instructors, not yet fully familiar with the system, did not insist that new speech patterns be collected when it was necessary. Perhaps these two students just happened to make a lot of stylization errors. It seems clear that good speech recognition accuracy is possible, at least for some talkers, since trainees S and B had recognition accuracy rates in excess of 90 percent.

TABLE 15. PERFORMANCE TEST SPEECH RECOGNITION ERRORS

Trainee	Phrase(s) Spoken	Phrase(s) Recognized
A	Turn right heading 155	Turn right heading 160
	Approaching glidepath	?
	5 miles from touchdown	?
	Slightly below glidepath	?
	Slightly below glidepath	Slightly above glidepath
	Slightly below glidepath	?
	Going further below glidepath	?
	Below glidepath	?
	Going further below glidepath	?
	Well below glidepath	?
	Going further below glidepath	?
	2 miles from touchdown	3 miles from touchdown
	Wind 180 at?	???? at ?
	Going above glidepath	?
	Above glidepath	?
R	Turn right heading 155	Turn right heading ?[1]
	5 miles from touchdown	?
	Turn right heading 162	Turn right heading ?[1]
	Slightly right of course	Slightly left of course[2]
	Correcting	?
	Slightly right of course	Slightly left of course
	Correcting	?
	Slightly right of course	Slightly left of course[2]
	Turn right heading 163	Turn right heading ?[1]
	Turn right heading 162	Turn right heading ?[1]
	Cough	Begin descent
	Cough	Wind
S	Coming down	Coming up
	1 mile from touchdown	2 miles from touchdown
	Button 1 clear	?
H	Marine 687	?
	Approaching glidepath	?
	Turn right heading 163	Turn right heading 165
	Coming up	Coming down
	Going below glidepath	Going above glidepatn[2]
	Correction	?
	3 miles from touchdown	?
	2 miles from touchdown	3 miles from touchdown
	Going above glidepath	Slightly above glidepath
	1 mile from touchdown	?
	Going below glidepath Slightly below glidepath At decision height	?[3]
	Coming up	?
	Slightly below glidepath	?

TABLE 15. PERFORMANCE TEST SPEECH RECOGNITION ERRORS (CONT)

Trainee	Phrase(s) Spoken	Phrase(s) Recognized
P	Coming up	Coming down
	4 miles from touchdown	2 miles from touchdown
	Going above glidepath	Slightly above glidepath
	Turn right heading 161	?[4]
	At decision height	?
	Over landing threshold Slightly right of course	?[5]
	Contact tower after landing	?
B	Turn right heading 155	Turn right heading 160
	Coming up	Coming down
	Turn left heading 160	Turn right heading 160
	Turn left heading 140	Turn right heading 140
	Too far right for safe approach	?
	1 mile	?

[1] Potentially a stylization error.

[2] SUS probably gave the trainee the benefit of the doubt when the trainee's phraseology was wrong.

[3] Counted as 3 unrecognized phrases, however the trainee did not pause between them.

[4] Counted as 4 unrecognized phrases, but the trainee did not pause between them.

[5] Counted as 2 unrecognized phrases, but the trainee did not pause between them.

Note: Only Trainees A and R finished the course during the training effectiveness testing period.

From a training perspective, these raw recognition data are not particularly informative because they do not show the significance of the recognition errors which occur. A recognition error is most significant if it causes the aircraft to behave in the wrong way. This not only appears very unrealistic, but can also produce a control situation which is much more difficult to handle. Table 17 shows the effect of the recognition errors on the aircraft dynamics. In this table, complete transmissions are described rather than isolated phrase elements as in the previous table. A complete transmission may consist of several phrases and constitutes a unit of control information. The table shows the total number of complete transmissions used by each trainee during the performance test, and the number of transmissions in which an error occurred which affected aircraft dynamics. Many recognition errors are simply not significant in terms of the actual dynamic situation because they do not involve control information. Thus the recognition error has no observable effect upon the simulated environment. Those errors which do affect the

TABLE 16. SUMMARY OF RECOGNITION RESULTS BASED UPON ANNOTATED PERFORMANCE TEST RESULTS

Trainee	Number of Phrases Spoken*	Number of Unrecognized Phrases	Number of Misrecognized Phrases	Number of Spurious Recognitions	Percent Unrecognized	Percent Misrecognized	Percentage Recognition Accuracy**
A	82	16	4	0	19.5%	4.9%	75.6%
R	100	15	3	2	15.0	3.0	82.0
S	86	2	2	0	2.3	2.3	95.4
H	97	10	5	0	10.3	5.2	84.5
P	77	8	3	0	10.4	3.9	85.7
B	98	2	5	0	2.0	5.1	92.9

* "Phrase" refers to a recognition unit as defined in Table 11.

** Percentage recognition accuracy is an overall measure which ignores the spurious recognitions.
It is computed as: $\frac{\text{Number of phrases correctly recognized}}{\text{Number of Phrases Spoken}} \times 100$

Note: Only trainees A and R participated during the actual training effectiveness testing period.

TABLE 17. EFFECT OF MISRECOGNITIONS AND UNRECOGNIZED PHRASES UPON AIRCRAFT DYNAMICS IN PERFORMANCE TEST RECORDS

Trainee	Number of Transmissions[1]	Number of transmissions in which errors occurred which affected aircraft dynamics[2]	Number resulting in no change in dynamics when change is expected[3]	Number resulting in a change in the right direction but in the wrong degree[4]	Number resulting in a change opposite to the desired effect
A	45	7	5	1	1
R	55	4	4	0	0
S	48	0	0	0	0
H	62	4	2	2	0
P	48	2	1	1	0
B	50	3	0	1	2

1 Defined as a complete transmission, e.g., "Marine 687, turn right heading 160, over."

2 Those which do not affect aircraft dynamics include pattern controller dialog, course position, glidepath trend, etc.

3 These transmissions were for the most part not recognized.

4 For example the digits in a turn may be taken from the model controller when a digit is misrecognized. Therefore sometimes the pilot understands "Turn right heading 160" when "Turn right heading 155" was spoken.

simulated environment are further categorized in the table. As might be expected, for trainees A and R for whom the rejection rate was extremely high, there is a proportionately large number of cases in which no change in aircraft dynamics occurred when such a change was expected on the basis of the transmission used. A "say again" feature in the pilot model might be helpful in cases like this for it would enable the system to notify the trainee that an unrecognized phrase should be repeated. This feature was not implemented because at system design time, and indeed until quite recently, the PAR controller was taught to keep the transmitter keyed continuously. This prevented the pilot from being able to transmit a "say again" request. Now that a change in training philosophy has occurred, the system could be modified to handle these unrecognized phrases more effectively.

The next column shows the number of recognition errors which caused a change in aircraft dynamics in the right direction but in the wrong degree. From a training perspective these aren't too serious. They include cases in which "going above glidepath" was recognized as "slightly above glidepath." They also include cases in which there was some problem in recognizing the digits in a turn command so the model controller's turn heading was chosen, resulting in a heading which was slightly different from the heading actually assigned. In the laboratory system evaluation this approach was suggested as being preferable to having the aircraft make no response. The method has its disadvantages in that it compensates for a trainee's poor control measures without his being aware that it is happening. However, at least in these cases, the frequency of this is so low that it is doubtful that the trainee would learn to rely on the system to correct his mistakes.

The last column in the table shows the most serious recognition errors. These actually cause the aircraft to maneuver in the direction opposite to the specified direction and can induce control situations which are difficult to handle. In the six performance test printouts which were made available to us, this happened 3 times out of the 308 transmissions uttered. This is rather impressive because it means that, for the most part, the recognition system is able to distinguish between the many similar phrases in the PAR controller vocabulary which cause opposite control effects. This was a matter of major concern in the design of the recognition algorithm.

What has not been completely analyzed is the effect of the recognition failures on performance measurement. The effect here is far less crucial to training effectiveness than the effect on aircraft dynamics. In fact, a feature of GCA-CTS permits the correction of recognition errors and subsequent automatic rescoring of the performance test. Still, recognition failures can cause poor scores when the real problem is poor recognition. The performance measurement system doesn't balk at all recognition errors of course, just as the aircraft dynamics aren't affected by them all. However, when the phrase "slightly above glidepath" is recognized and "going above glidepath" was actually spoken, a trend error will be counted. Likewise, when the trainee says "turn left" and the system hears "turn right," an error is also counted. Since performance test scores can be modified automatically to correct for misrecognitions, the scores before and after modification can be compared. Table 18 shows these data for the four students for whom they are available. The system always explains the scoring rationale to the student during annotated replay by reporting, "you were understood to say:..." then describing the error and giving the correct phraseology. Therefore the reason for a poor score due to a recognition error is evident to the trainee. Still, these poor scores don't reflect his actual performance and therefore are inaccurate. Or are they, really? What is at the bottom of it? Some trainees can get excellent recognition, as shown in Table 16. It would be interesting to know what is really going on. Is the poor recognition due to sloppy speech? For example, do some trainees tend to swallow the first part of the utterance such that there is very little difference between "going above glidepath" and "slightly above glidepath"? If so, this lazy voicing ought to be discouraged with low performance measurement scores. A real pilot listening to a radio may not hear the trainee any better.

Then again, is the poor recognition due to stress-induced pitch changes in the trainee's voice? If so, perhaps the trainee should be encouraged to reevaluate how well suited he or she is for air traffic control responsibilities. Are low scores really undesirable in such a case? Perhaps not if a potential reason for them is recognized to be undue reaction to stress. Research may reveal that poor recognition has potential as a selection criterion.

Is the poor recognition due to the courseware? Could the recognition errors be further decreased by amplifying the courseware in the later levels to better elicit natural speech patterns? An interesting observation has been made with respect to this, specifically that speech recognition doesn't work as well during a problem as it does in the validation mode, especially in levels 4 and above. There are a number of possible reasons for this, and some of them have been investigated. It may be that the graphics display processing is stealing too much CPU time and that input feature patterns are being lost during a problem. This has not been tested, but the observation that experienced talkers get good recognition results in any level suggests that this cannot be having a major impact if it is occurring at all. To test the hypothesis that perhaps the model controller algorithm was getting behind and biasing the recognition algorithm in an undesirable way, this feature was disabled, but the recognition did not seem to improve. If it were simply the case that the recognition algorithm can't handle the larger vocabulary which is used in level 4 and above, the validation mode would show a similar decline in recognition accuracy. There is, however, one big difference between level 4 and the previous levels, and that is that it provides minimal instruction and asks the trainee to repeat phrases out of context for voice data collection. (The reasons for this were explained in the courseware development subsection.) Could it be then, that the decline in recognition accuracy is due, at least to some extent, to the tendency to say things differently when in the midst of a control situation then when repeating the phrase out of context? It was in full awareness of this tendency that the system was designed to have the capability to provide the kind of instruction implemented

TABLE 18. EFFECT OF RECOGNITION ERRORS ON PERFORMANCE MEASUREMENT SCORES

Trainee	S		H		P		B	
Skill Category	Original	Modified	Original	Modified	Original	Modified	Original	Modified
Handoff acceptance	85		100		100		100	
Radio check	100		100		100		100	
Turn to final	100		80	70	100		100	
Approaching glidepath	95		95	100	80		90	
Heading transmissions	100		100		86	90	92	100
Azimuth position and trend	100		100		100		100	
Glidepath position and trend	93	95	88	96	94		94	
Range calls	96	100	60	96	80		100	
Decision height	100		50	100	50	75	50	25
Clearance	100		100		100		100	
Over landing threshold	40		40		60	80	100	
Handoff or rollout	60		80		60		90	
Emergency waveoff	100		100		100		25	
Transmission break	100		100		80		100	
Transmission rate	100		90	93	100		100	

in levels 2 and 3. An experiment could easily be designed to test the hypothesis that there is no difference in recognition accuracy attributable to the different kinds of training.

Finally, could poor recognition be due to failure to conform to the required stylization? Perhaps trainees tend to forget the necessary stylization when they have to think about the control situation. There are cases in the performance test printouts which show that the trainee did not pause between phrases. Further, observation indicated that trainees commonly inserted pauses in the midst of a phrase. Either stylization error will prevent recognition. With an isolated phrase recognition system there is very little that can be done about the former stylization error. The speech understanding algorithm in GCA-CTS does attempt to concatenate some fragmentary utterances, however the trainees made up some unexpected ones! There is a limit to how much of this concatenation can be done because each fragment adds another potential recognition to the already extensive vocabulary list. It should be noted that the replacement of the isolated phrase recognition device with a continuous speech recognition device will not automatically solve this latter stylization problem. It must be taken into consideration with any type of recognition technology.

SYSTEM RELIABILITY. The reliability data obtained during the actual training effectiveness test period were low for the reasons described in the subsection on the training system package. Reliability improved dramatically after the problems described here were corrected.

SECTION IV

RECOMMENDATIONS

In this section, some problem areas are noted and some recommendations are made regarding a possible production system. Concluding remarks complete the report.

PROBLEM AREAS

Some problems encountered in the development of GCA-CTS are discussed in this subsection. The focus is not upon the specific difficulties encountered in design or implementation — these have been described in previous paragraphs. Ths focus here is upon some more general problems which might be encountered in other development efforts of this type.

COURSEWARE VALIDITY. In the development of the behavioral objectives and courseware, the problem of validating the material was continually present. First of all, although the subject matter experts were extremely helpful, courteous and patient, as outsiders trying to learn the details of the GCA control task we encountered some typical communication difficulties. One form arose as "informant bias" which Diesing (1971) has summarized well as follows:

> Informants' statements are biased in various ways, depending on the topic, the circumstance, the informant, and his relationship to the researcher. To a newcomer an informant is likely to give an idealized version of what happens or how he feels; a co-operative informant may report what he thinks the researcher wants to hear or would find interesting; esoteric material is likely to be simplified to the researcher's presumed level of comprehension; ...perceptual and cognitive distortions may occur in material in which the informant is personally involved and important details may be missing in material in which he is not involved; and so on.

These problems were resolved by cross-checking all of the information with at least two or three subject matter experts to achieve contextual validity.

Another communications difficulty arose from the difference in perspectives held by the subject matter experts on the one hand and the analysts on the other. In order to build a computer model of controller behavior, it is necessary to follow out the ramifications of each aspect of the behavior systematically and in excruciating detail. Quantification must replace "gut-feel." However, this perspective was quickly and graciously adopted by the subject matter experts with whom we worked and their willingness to learn our language proved immensely helpful.

Finally communications were made difficult by our lack of proximity to the experts, and to the continually developing nature of the task itself: regulations change, stylistic changes occur as personnel changes take place, equipment is modified, etc. At times even our questions, when systematically pursued, would lead to procedural changes in the interest of consistency. It was difficult to keep abreast of these changes. Some of the behavioral objectives which were validated in early 1978 were no longer valid in late 1979 at system delivery time. The training system was, however, designed to be flexible enough that necessary modifications could be made. Indeed, changes were made during system delivery and further enhancements were added later.

To summarize then, even though we were aware of the potential communications pitfalls and have endeavored to avoid them, we still found it necessary to fall back on the system's inherent ability to be changed. In future developments, both maintenance of communications and flexible system design should continue to be stressed.

DOCUMENTATION OF RESULTS. The GCA-CTS is being studied intensively and much of this work is, or will be, documented. It is possible, however, that the real experts, namely the learning supervisors who have devoted an immense amount of interest, time and patience to the system will not be sufficiently debriefed. These are the only people in the world who have ever tried to use an instructorless training system based on the automated speech technologies. It would be a great loss if their observations, ideas and suggestions were not adequately elicited and preserved.

GOVERNMENT-FURNISHED EQUIPMENT. Several problems were encountered due to the Government-furnished equipment. Local hardware people discovered that the processors were different in some respects from those sold on the west coast. It is conceivable that the availability of some components might have been a problem, and is something to consider in future procurements.

SYSTEM DEVELOPMENT HARDWARE. A woe is expressed here for which there may be no adequate solution. It is this: it is difficult to develop a system on the hardware for which it is intended. Of course there are many advantages which must not be discounted, such as the lack of interoperating system incompatibilities, and so on. But there are significant problems as well. With only two CRTs, only a maximum of two persons could use the GCA-CTS system at a time. During integration, it would support only one person. If two were using it, they had to share disk resources. This was a problem not only because there was only one removable cartridge and very little room on the fixed disk for user subdirectories, but also because such disk I/O bound processes as Fortran compiles would slow response time to the other user. Only one user had direct access to the printer. The printer, though ideal for instructional system purposes, was inadequate for program development. The lack of a magnetic tape system caused untold grief ranging from dependence upon computer center time availability for disk backups, to aching backs from lugging five disks to Memphis when a tape or two would have carried the same data (and we might add, more safely). These minimal system resources were in use virtually 24 hours per day throughout the development effort. It would be interesting to examine the tradeoffs systematically and to determine whether increased programmer productivity might be realized if some concessions to the development effort were made in terms of hardware resources. Some compromise between minimum system hardware requirements and those which would ideally support system development might even prove to be cost effective.

PRODUCTION SYSTEM

The remarks in this section are addressed to the development of an engineering prototype system from the experimental prototype baseline system. The existing GCA-CTS supports only one trainee station. Presumably a production system would consist of several trainee stations, and the learning supervisor station would have a graphics capability such that any trainee station could be monitored. The multistation concept would require a new design effort in which the existing system would be treated as a sort of predesigned module to be attached to a larger whole. Some ramifications of this concept are described below.

HARDWARE. Additional disk and computational resources would be needed to support a multistation GCA-CTS. Prior to procurement of a speech recognition system, however, the need for and feasibility of a continuous speech recognition capability must be assessed.

The whole realm of face validity also must be considered. The need to provide sweep and the means to do it must be determined. The need for a look-alike radar console, and the ability to simulate radar clutter, gain control effects, etc., should be addressed. The need for a better CRT positioning scheme has already been mentioned.

Small changes are also needed to trainee panel design to enable the learning supervisors to plug their headsets into the trainee station, in part because the speaker has been found to disrupt speech recognition.

SOFTWARE. As with any new software, there are modules within GCA-CTS which could benefit from improvements. Many such suggestions have been described in previous paragraphs. For example, an obvious improvement would be to provide the controller model and pilot with digitized speech voices. With additional disk space, this could be done.

INSTRUCTION. Many instructional questions have been raised in the previous paragraphs and these point to the need for the courseware to be validated, as all courseware must be. This must be done at two levels. First, the question of which approach is better, that taken in levels 2 and 3 or that taken in the remainder of the syllabus, must be answered. At the next level, the specific courseware materials must be validated. For example, the instructors have made numerous suggestions for improvement of the textual materials, and these suggestions need to be taken into consideration.

CONCLUDING REMARKS

This report chronicles a project which has been challenging because of the great many problems to be solved and new ideas to be implemented. It is rewarding to note that it is enjoying at least modest success and especially that it has stimulated so much thinking about how best to accomplish the automated training of a verbal task. It is hoped that this document, in describing the rationale behind the design decisions that were made, will provide an important reference for future researchers who are called upon to evaluate the results of these decisions and to suggest improved methodologies. Thus the experimental prototype GCA-CTS and its documentation will have served its purpose: to provide one more step toward the goal of applying the automated speech technologies to training in an effective way.

REFERENCES

Barber, G.D., J. Bollenbacher, D. Brewton, D. Harry, M. Hicklin, C. Meyn, G.K. Slemon, Ground Controlled Apprcach Controller Training System System Documentation, NAVTRAEQUIPCEN 77-C-0162-3, 1980.

Barber, G.D., and M. Hicklin, Precision Approach Radar Controller Training System Functional Specification, NAVTRAEQUIPCEN 77-C-0162-7, 1980.

Barber, G.D., M. Hicklin, C. Meyn, J.E. Porter, G.K. Slemon, Ground Controlled Approach Controller Training System Design Report, NAVTRAEQUIPCEN 77-C-0162-2, 1978.

Breaux, R., Training Characteristics of the Automated Adaptive Ground Controlled Approach Radar Controller Training System (GCA-CTS), Technical Note TN-52, July 1976.

Breaux, R. and M.W. Grady, "The Voice Data Collection Program. A Generalized Tool for Studies in Speech Recognition," in Ninth NTEC/Industry Conference Proceedings, Technical Report NAVTRAEQUIPCEN IH-276, 1976.

Diesing, P., Patterns of Discovery in The Social Sciences, Aldine Publishing Company, 1971.

Feuge, R.L., J.P. Charles and R. Miller, Feasibility of Automated Adaptive GCA Controller Training System, Technical Report: NAVTRAEQUIPCEN 73-C-0079-1, April 1974.

Goldstein, I., D.A. Norman, et al., "Ears for Automated Instruction Systems: Why Try?" in Proceedings of the Seventh NAVTRAEQUIPCEN/Industry Conference, NAVTRAEQUIPCEN IH-240, 1974.

Grady, M.W. and M. Hicklin, Use of Computer Speech Understanding in Training: A Demonstration Training System for the Ground Controlled Approach Controller, Technical Report: NAVTRAEQUIPCEN 74-C-0048-1, July 1976.

Hicklin, M., L.H. Nowell and R. Petersen, Ground Controlled Approach Controller Training System Training/Functional Design Report, NAVTRAEQUIPCEN 77-C-0162-1, 1978.

Hicklin, M. and G. Slemon, "User Acceptance In an Automated Speech Technology Based Training System," in Eleventh NTEC/Industry Conference Proceedings, Technical Report NAVTRAEQUIPCEN IH-306, 1978.

Hicklin, M. et al., Ground Controlled Approach Controller Training System Instructor Guide, NAVTRAEQUIPCEN 77-C-0162-5, 1980a.

Hicklin, M. et al., Ground Controlled Approach Controller Training System Student Guide, NAVTRAEQUIPCEN 77-C-0162-4, 1980b.

Logicon, Inc., Tactical and Training Systems Division, Ground Controlled Approach Controller Training System Demonstration Test Plan, 1979a.

Logicon, Inc., Tactical and Training Systems Division, Ground Controlled Approach Controller Training System Demonstration Test Report, 1979b.

APPENDIX A

DATA DELIVERED

Table A-1 shows the data delivered during the course of the GCA-CTS project and provides a brief statement of the purpose for each item.

TABLE A-1. DATA DELIVERED

Item	Report Number (77-C-0162-)	Purpose
Work Plan Report	-	Described project schedule
Quarterly Progress Reports	-	Described progress of project
Training/Functional Design Report	1	Listed behavioral objectives, described the syllabus and system functional requirements
System Configuration/ Facilities Report	2	Specified GCA-CTS design, also on-site facilities requirements
Demonstration Test Plan	-	Described the GCA-CTS acceptance tests
Trainer Demonstration Report	-	Described the results of the demonstration tests
Training Effectiveness Test Plan	-	Described a training effectiveness testing scheme
Training System Package		Included the following documentation:
	-	Commercial hardware documentation
	5	Instructor Guide
	4	Student Guide
	3	System Documentation
	-	Program listings
	-	Courseware listings
Final Technical Report	6	The present document
Functional Specification	7	A functional specification for the retrofit of the existing training device with a GCA-CTS capability.

ACRONYMS

APE	Aircraft/Pilot/Environment
CAI	Computer-Assisted Instruction
GCA-CTS	Ground Controlled Approach-Controller Training System
GFE	Government-Furnished Equipment
IPB	Interprocessor Bus
NATTC	Naval Air Technical Training Center
NAVAIRSYSCOM	Naval Air Systems Command
NAVEDPRODEVCEN	Naval Education and Training Program Development Center
NAVTRAEQUIPCEN	Naval Training Equipment Center
PAR	Precision Approach Radar
PMS	Performance Measurement Subsystem
VDC	Voice Data Collection
VTAG	Compromise Acronym of the Voice Technical Advisory Group

DISTRIBUTION LIST

Naval Training Equipment Center 97
Orlando, FL 32813

Defense Documentation Center 12
Cameron Station
Alexandria, VA 22314

Army Training Support Center 2
ATTSC-DE
Ft. Eustis, VA 23604

All other addressees receive one copy of report

National Aviation Facilities
Experimental Center
Library
Atlantic City, NJ 08405

Dr. J. D. Fletcher
Defense Adv. Research Projects
Agency (CTO)
1400 Wilson Blvd.
Arlington, VA 22209

Dr. Donald Connolly
Federal Aviation Administration
FAA NAFEC ANA-230
Atlantic City, NJ 08405

Chief, Methodology & Standards Staff
FAA Academy
Aeronautical Center, AAC - 914
P.O. Box 25082
Oklahoma City, OK 73125

Dr. John F. Boehm
Director, National Security Agency
9800 Savage Road
Attn: R-542 Boehm
Ft. George G. Meade, MD 20755

Dr. Jesse Orlansky
Institute for Defense Analyses
Science and Technology Division
400 Army-Navy Drive
Arlington, VA 22202

John Sigona
Department of Transportation
DOT/TSC-533
Kendal Square
Cambridge, MA 02142

Mr. Harold C. Glass
U.S. Postal Lab
11711 Parklawn Drive
Rockville, MD 20852

Mr. John T. Masterson
U.S. Postal Lab
11711 Parklawn Drive
Rockville, MD 20852

Captain Vince Mortimer
AFAMRL/BBM
Wright-Patterson AFB
Dayton, OH 45433

Mr. Eric Werkowitz
AFFDL/FGR
Wright-Patterson AFB
Dayton, OH 45433

Mr. R. G. Cameron
ASD/ENETC
Wright-Patterson AFB
Dayton, OH 45433

Mr. Thomas J. Moore
AFAMRL/BBM
Wright-Patterson AFB
Dayton, OH 45433

Mr. Don F. McKechnie
AFAMRL/HEF
Wright-Patterson AFB
Dayton, OH 45433

Mr. Harold A. Kottmann
ASD/YWE
Wright-Patterson AFB
Dayton, OH 45433

U.S. Air Force Human Resources Lab
AFHRL/OTR
Flying Training Division
Williams AFB, AZ 85224

Air Force Human Resources Lab
AFHRL/LR
Logistics Research Division
Wright-Patterson AFB
Dayton, OH 45433

U.S. Air Force Human Resources Lab/MF
Brooks AFB, TX 78235

U.S. Air Force Human Resources Lab
AFHRL-MPM
Manpower & Personnel Division
Manpower & Force Management Systems Br.
Brooks AFB, TX 78235

U.S. Air Force Human Resources Lab/TSZ
Brooks AFB, TX 78235

U.S. Air Force Human Resources Lab
AFHRL-TT (Dr. Rockway)
Technical Training Division
Lowry AFB, CO 80230

Director
Air University Library
Maxwell AFB, AL 36100

Headquarters 34 Tactical Airlift
Training Group/TTDI
Little Rock AFB, AK 72076

Commanding Officer
Rome Air Development Center Library
TSLD
Griffiss AFB, NY 13441

LT Jeff Woodard
RADC/IRAA
Griffiss AFB, NY 13441

Headquarters
Air Training Command, XPTI
Attn: Mr. Goldman
Randolph AFB, TX 78148

Dr. Tice DeYoung
U.S. Army Engineer Topographic
Laboratories Research Institute
Ft. Belvoir, VA 22060

Dr. Henry Dehaan
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Capt. Leslie K. Scofield
Directorate of Training
U.S. Army Signal Center
Fort Gordon, GA 30905

Director, U.S. Army Research and
Technology Laboratory
DAVDL-AS (Attn: Dr. R. S. Dunn)
207-S Ames Research Center
Moffett Field, CA 94035

USAHEL/USAAVNC
Attn: DRXHE-FR (Dr. Hofmann)
P.O. Box 476
Ft. Rucker, AL 36362

LTCOL Ray Kreulen
USAAVNC, DCAT-DT (ATZO-T-TM)
P.O. Box 395
Ft. Rucker, AL 36362

Mr. Lockwood Reed
U.S. Army Avionics R&D Activity
DAVAA-E
Ft. Monmouth, NJ 07703

Commandant
U.S. Army Field Artillery School
ATSF-TD-TS (Mr. Inman)
Ft. Sill, OK 73503

Commandant
U.S. Army Field Artillery School
Counterfire Department (Mr. Rogers)
Ft. Sill, OK 73503

Mr. Clayton Coler
NASA AMFS Research Center
Mail Stop 239-2
Moffett Field, CA 94035

Mr. Charles E. Lueck, Jr.
U.S. Postal Service
Process Control Systems Test Facility
9201 Edgeworth Drive
Washington, D. C. 20027

Commandant of Marine Corps
Code OTTF 31
Washington, D. C. 20380

Scientific Advisor
Headquarters, U. S. Marine Corps
Washington, D. C. 20380

Mr. Gary Poock
Naval PG School
Code 55 PK
Monterey, CA 93940

Commander, Naval Air Systems Command
AIR 340F
Washington, D. C. 20361

Commander, Naval Air Systems Command
AIR 413
Washington, D. C. 20361

Commander, Naval Air Systems Command
AIR 413E (Mr. Ernest E. Poor)
Washington, D. C. 20361

Commander, Naval Air Systems Command
AIR 413F
Washington, D. C. 20361

Mr. John Martins, Jr.
Naval Underwater Systems Center
New London Laboratory MC 315
New London, CT 06320

Commander, Naval Air Systems Command
Washington, D. C. 20361

Commander
Naval Weapons Center
Code 3154 (Mr. Bob Curtis)
China Lake, CA 93555

Commander
Naval Air Development Center
Code 4043 (T. Weiner)
Warminster, PA 18974

Commander
Naval Air Development Center
Code 6021 (Christian P. Skriver)
Warminster, PA 18974

Commander
Naval Air Development Center
Code 6022
Warminster, PA 18974

Chief of Naval Material
MAT 08D2 (A. Rubinstein)
CP5, Room 678
Washington, D. C. 20360

CDR Richard S. Gibson
Bureau of Medicine and Surgery
Head, Aerospace Psychology Branch
Code 3C13
Washington, D. C. 20372

John Silva
Naval Ocean Systems Center
Code 823
San Diego, CA 92152

Mr. Harry A. Whitted
Naval Ocean Systems Center
Code 8235
San Diego, CA 92152

Dr. Sam Schiflett
Naval Air Test Center
SY 721
Patuxent River, MD 20670

Dr. Roger Remington
Naval Aerospace Medical Research Lab
Code L52
Pensacola, FL 32508

Commanding Officer
Naval Air Technical Training Center
Code 104, Building S-54
NAS Memphis (85)
Millington, TN 38054

Commanding Officer
Naval Air Technical Training Center
Code 7411
NAS Memphis (85)
Millington, TN 38054

Commander
Naval Air Development Center
Code 602 (CDR P. M. Curran)
Warminster, PA 18974

Commander
Naval Air Development Center
Code 6021 (CDR Norman E. Lane)
Warminster, PA 18974

Commander
Naval Air Development Center
Code 6021 (LT Steve Harris)
Warminster, PA 18974

Commanding Officer
Fleet Combat Training Center, Pacific
Code 09A
San Diego, CA 92147

Commander, Naval Air Force
U.S. Pacific Fleet (Code 316)
NAS North Island
San Diego, CA 92135

Commander, Naval Air Force
U.S. Pacific Fleet (Code 342)
NAS North Island
San Diego, CA 92135

Navy Personnel Research & Development Center
Attn: M. McDowell
Library, Code P201L
San Diego, CA 92152

Navy Personnel Research & Development Center
Attn: Dr. Robert Wisher, Code P309
San Diego, CA 92152

Navy Personnel Research & Development Center
Attn: Mr. R. Obermayer, Code P31
San Diego, CA 92152

Chief of Naval Operations
Code 221 (Mr. J. Trimble)
800 N. Quincy St.
Arlington, VA 22217

Chief of Naval Research
Code 458 (Mr. Henry Halff)
800 N. Quincy St.
Arlington, VA 22217

Chief of Naval Research
Code 458 (Dr. Marshall Farr)
800 N. Quincy St.
Arlington, VA 22217

Chief of Naval Operations
Code OP-39T (Mr. Walt Primas)
Washington, D. C. 20350

Chief of Naval Operations
OP-513
Washington, D. C. 20350

Chief of Naval Operations
OP-596C
Washington, D. C. 20350

Chief of Naval Operations
OP-987H (Dr. R. G. Smith)
Washington, D. C. 20350

Deputy Chief of Naval Operations (MPT)
OP-102
Arlington Annex
Washington, D. C. 20350

Commander
Naval Sea Systems Command
Code 6122 (Mr. P. J. Andrews)
Washington, D. C. 20362

Dr. James Mosko
Naval Aerospace Medical Research Lab
Acoustical Sciences Division
Code L340
Pensacola, FL 32508