

AD-A035 637

FRANKFORD ARSENAL PHILADELPHIA PA
VOICE OUTPUT/INPUT CONTROL OF EQUIPMENT (VOICE).(U)
OCT 75 J F HALGAS
FA-FCF-3-76

F/G 17/2

UNCLASSIFIED

NL

1 of 1
ADA035637



END
DATE
FILMED
3-77

ADA 035637

J. Halgas 12

FCF-3-76

FINAL REPORT

VOICE OUTPUT INPUT CONTROL OF EQUIPMENT
(VOICE)

BY
JOSEPH F. HALGAS

OCTOBER 1975

DDDC
FEB 15 1977
REGISTRATION
QC



FIRE CONTROL DEVELOPMENT
AND ENGINEERING DIRECTORATE

U.S. ARMY ARMAMENT COMMAND
FRANKFORD ARSENAL
PHILADELPHIA, PENNSYLVANIA 19137

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

FINAL REPORT
VOICE OUTPUT INPUT CONTROL OF EQUIPMENT
(VOICE)

by

JOSEPH F. HALGAS

| | |
|------------------------------------|---|
| ACCESSION FOR | |
| NTIS | White Section <input checked="" type="checkbox"/> |
| U. C. | Buff Section <input type="checkbox"/> |
| DWARDED BY | |
| JUSTIFICATION | |
| BY DISTRIBUTION/AVAILABILITY CODES | |
| Dist. | APPL. AND/OR SPECIAL |
| A | |

OCT 1975

U.S. ARMY ARMAMENT COMMAND
FRANKFORD ARSENAL
PHILADELPHIA, PENNSYLVANIA 19137

REPORT #FCF-3-76

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|--|--|
| 1. REPORT NUMBER FA-FCF-3-76 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) Voice Output/Input (VOICE) Control of Equipment (VOICE) | 5. TYPE OF REPORT & PERIOD COVERED Final Report | 6. PERFORMING ORG. REPORT NUMBER |
| | | 8. CONTRACT OR GRANT NUMBER(s) |
| 7. AUTHOR(s) Joseph F. Halgas | 9. PERFORMING ORGANIZATION NAME AND ADDRESS Frankford Arsenal ATTN: SARFA-FCF-I Philadelphia, Pa. 19136 | |
| 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element: 6-36-22.A Project: IG663622AJ29 | | 11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Armament Command ATTN: AMSAR-RDS Rock Island, Ill. |
| 12. REPORT DATE Oct 1975 | | 13. NUMBER OF PAGES 33 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Frankford Arsenal ATTN: SARFA-FCF-I Philadelphia, PA 19137 | | 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) VOICE OUTPUT/INPUT CONTROL OF EQUIPMENT (VOICE) AUTOMATIC TEST, MEASUREMENT AND DIAGNOSTIC EQUIPMENT (ATMDE) MAN-MACHINE INTERFACE (MMI) UNIT UNDER TEST (UUT) | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This is a report of an investigation of Voice Output/Input equipment as a practical Output/Input subsystem for Automatic Test, Measurement and Diagnostic Equipment and other computer based operations. The feasibility of replacing present modes of man-machine interface by Voice Output/Input Control of Equipment (VOICE) is demonstrated through use of commercially available peripherals configured into a closed loop Voice O/I system. The Closed Loop Voice O/I system is based on isolated word and phrase (discrete) | | |

142 720 KB

processing which is fully compatible with formats employed in most military functions. The report includes descriptions of presently available Voice Output/Input equipment, the state-of-the-art system configured, performance characteristics and potential areas of application.

19. KEY WORDS (continued)

Synthesized Voice Response Equipment (SVRE)

Voice Recognition Equipment (VRE)

Fault Isolation

Fault Location

Closed Loop Voice Output/Input

Real Time Processing

Discrete Speech Recognition

Continuous Speech Recognition

Adaptive Speech Recognition

Universal Speech Recognition

Contextual Speech Response

Non-Contextual Speech Response

TABLE OF CONTENTS

| <u>SECTION</u> | | <u>PAGE</u> |
|----------------|--|-------------|
| I | <u>INTRODUCTION</u> | |
| | 1.2 AUTOMATION OF TEST, MEASUREMENT AND DIAGNOSTIC EQUIPMENT | 1 |
| | 1.2 THE MAN-MACHINE INTERFACE | 1 |
| II | <u>VOICE OUTPUT/INPUT CONTROL OF EQUIPMENT (VOICE)</u> | |
| | 2.1 VOICE OUTPUT/INPUT AS A VIABLE ALTERNATIVE TO PRESENT MODES OF MAN-MACHINE INTERACTION | 2 |
| | 2.2 OBJECTIVES | 2 |
| III | <u>STATE-OF-THE-ART VOICE SYSTEM</u> | |
| | 3.1 ESTABLISHING CRITERIA FOR VOICE RECOGNITION VOICE RESPONSE EQUIPMENT | 3 |
| | 3.2 VOICE RECOGNITION EQUIPMENT - VIP-100 | 4 |
| | 3.3 VOICE RESPONSE EQUIPMENT - S-NOVA | 6 |
| | 3.4 CLOSED LOOP VOICE SYSTEM | 8 |
| | 3.5 ASSOCIATED PERIPHERALS | 10 |
| IV | <u>VOICE SYSTEM EVALUATION</u> | |
| | 4.1 SYSTEM EVALUATION AND DEVELOPMENT | 15 |
| | 4.2 VOICE FEASIBILITY DEMONSTRATIONS | 18 |
| | 4.3 GENERAL PROGRAM FLOW OF DEMONSTRATIONS | 26 |
| | 4.4 NEW DEVELOPMENTS IN VOICE O/I | 26 |
| V | <u>CONCLUSIONS AND POTENTIAL APPLICATIONS</u> | |
| | 5.1 CONCLUSIONS | 32 |
| | 5.2 POTENTIAL APPLICATIONS | 33 |

LIST OF FIGURES

| <u>FIGURE</u> | | <u>PAGE</u> |
|---------------|--|-------------|
| 1 | VOICE RECOGNITION EQUIPMENT | 5 |
| 2 | VOICE RECOGNITION WORD PATTERN GENERATION | 7 |
| 3 | SYNTHESIZED VOICE RESPONSE EQUIPMENT | 7 |
| 4 | VOICE RESPONSE SYNTHESIZER DATA GENERATION | 9 |
| 5 | CLOSED LOOP VOICE O/I PROGRAM | 11 |
| 6 | STATE-OF-THE-ART VOICE SYSTEM | 13 |
| 7 | VOICE OPERATING SYSTEM CONSOLE | 14 |
| 8 | AUDIO AMPLIFIER CIRCUIT | 20 |
| 9 | M151 ENGINE ILLUSTRATION | 23 |
| 10 | GENERAL PROGRAM FLOW OF DEMONSTRATIONS | 27 |

LIST OF TABLES

| <u>TABLE NO.</u> | | <u>PAGE</u> |
|------------------|---|-------------|
| I | PHONEME LIST | 17 |
| II | AUDIO AMPLIFIER CIRCUIT TEST | 19 |
| III | M151 ENGINE DIAGNOSTIC PROCEDURE | 22 |
| IV | FIRE MISSION-FORWARD OBSERVER/FIRE DIRECTION CENTER | 24 |
| V | SAMPLE CALCULATION | 25 |

ACKNOWLEDGEMENTS

The system described in this report has been influenced by the interactions that the author had with individuals active in the field of Speech Recognition and Synthesized Speech Response.

Mr. E. Carroll, Frankford Arsenal, was primarily responsible for development of the software necessary to implementation of the various feasibility demonstrations. Other contributing personnel include Mr. G. Fenerty, a Drexel University Engineering student trainee, who assisted in the software development and hardware fabrication of VOICE and Miss J. Herman, also a Drexel Engineering student trainee, who provided modification of the VIP-100 and S-Nova support programs to accomplish closed loop operation and composition of a synthesized speech vocabulary.

The author has also benefited from discussions with Mr. Robert Cox, Threshold Technology, Inc. and Messrs. Robert Grisdale and Charles Peacock, Interface Systems, Inc.

SECTION I

INTRODUCTION

1.1 AUTOMATION OF TEST, MEASUREMENT AND DIAGNOSTIC EQUIPMENT

Increased complexity of modern weapon systems and the limited availability of highly skilled technical personnel in the military services has led to increased materiel maintenance costs. Currently, maintenance costs, during the systems life cycle, exceed the acquisition cost by a factor of ten (10). Major cost items in a maintenance program are the training of personnel and man-hours expended in testing and diagnosis of a Unit Under Test (UUT). Automatic Test, Measurement and Diagnostic Equipment (ATMDE) should, therefore, be designed to alleviate the test, measurement and diagnostic procedures required to locate and isolate the fault(s) in a UUT. Properly designed ATMDE is applicable not only to product maintenance but, in fact, can encompass the design, production and quality assurance stages of procured materiel.

While automation of maintenance procedures can effectively reduce maintenance costs, further savings can be realized by resolving the Man-Machine interface problem.

1.2 THE MAN-MACHINE INTERFACE

The Man-Machine Interface (MMI) via keyboards, punched cards or tape has always been, at best, an indirect mode of interaction. From an operations standpoint these modes of interaction tend to be very error prone. Exchange of data between operator and computer by means of a keyboard is a tedious, boring, unreliable and inefficient means of communication.

The most natural way for human beings to communicate is through the use of speech. This effort represents the investigations of Voice Output Input Control of Equipment (VOICE) as a practical output/input subsystem for ATMDE and other computer based operations.

SECTION II

VOICE OUTPUT INPUT CONTROL OF EQUIPMENT (VOICE)

2.1 VOICE OUTPUT/INPUT AS A VIABLE ALTERNATIVE TO PRESENT MODES OF MAN-MACHINE INTERACTION

Why Voice Output/Input? In preference to modes of man-machine interaction presently employed, voice O/I can offer many distinct advantages. Voice O/I minimizes personnel training requirements. Relatively unskilled personnel could use voice O/I controlled equipment. With voice O/I, it becomes possible for man-machine to interact directly when inputs indicate a need for special instruction. The interaction occurs in real-time which is a desired feature in applications requiring real time processing. Real time feedback to the operator, visual and/or verbal, can efficiently guide the operator through a complete process without interrupting the operator's motor activities. Voice output/Input yields additional benefits in the form of improved efficiency, cost reduction and increased reliability of the data collected and subsequently used in a given operation.

2.2 OBJECTIVES

The program was initiated with the objective of identifying, acquiring and integrating commercially available equipment into a functional VOICE sub-system as quickly as possible and at a minimum cost. Toward this end the following tasks were pursued:

- A. Technical studies and survey of Industry as to the availability of voice recognition and voice response equipment.
- B. Determination and acquisition of voice recognition and voice response equipment that most nearly meets criteria established by task A.
- C. Integration of voice recognition and voice response equipment into a closed loop VOICE system.
- D. Addition of visual display capability as an operator aid to increase the reliability of the information exchange process.
- E. Add hard copy capability for the purpose of recording data.
- F. Perform an experimental evaluation of the VOICE system configured. Demonstrate the feasibility of VOICE.
- G. Study and continued investigation of pertinent developments (other systems, new developments, techniques, etc.).

SECTION III

STATE-OF-THE-ART VOICE SYSTEM

3.1 ESTABLISHING CRITERIA FOR VOICE RECOGNITION-VOICE RESPONSE EQUIPMENT

To gain insight into the problems of voice recognition-voice response, the author conducted a literature study of Speech Analysis, Synthesis and Perception.^{1,2} Industrial concerns working in the field were surveyed and their capabilities appraised.

Criteria established for the Voice recognition equipment was as follows:

1. The voice recognition equipment was to be configured around a mini Computer. Processed acoustical features would be fed to the computer to be compared and classified with previously stored reference patterns for a given vocabulary of words. Through this technique real-time processing could be readily achieved.
2. Successful operation of the voice recognition equipment in a wide noise-background environment.
3. While continuous (connected) speech recognition was desired over discrete (isolated) speech recognition, the latter was directed due to the non-existence of any commercially available connected speech handling equipment.
4. The equipment must have an interactive feature that ensures that voiced inputs have been correctly interpreted by the machine.
5. Abnormalities such as head colds, sore throats or hoarseness should not hamper the recognition process.
6. Adaptive recognition equipment (equipment that can adapt to speech characteristics of the User) was specified over "Universal" recognition equipment (Equipment that can recognize a set vocabulary spoken by a wide range of users) to insure maximum flexibility and recognition accuracy.
7. Speech recognition equipment should feature ease of operation and a flexible vocabulary. Operating software should include a training mode, interactive recognition phase, teletype O/I control and a diagnostic program.

-
1. J. L. Flanagan, Speech Analysis, Synthesis and Perception. New York: Academic Press, 1965.
 2. R. K. Potter, D. A. Kopp and H. G. Kopp, Visible Speech. New York: Dover Publications, 1966.

Criteria established for the Voice response equipment was as follows:

1. The voice response equipment was to be based on the same mini-computer as the voice recognition equipment if at all possible. Response data must require a minimal amount of memory locations for each word response.
2. The equipment must have real-time capabilities.
3. Synthesized Voice response must be highly intelligible, Contextual as well as non-contextual capability must be featured.
4. Speech response equipment should feature ease of operation and a flexible vocabulary. Operating software should include a composer (word and message) mode, response phase, teletype O/I control and a diagnostic program.

3.2 VOICE RECOGNITION EQUIPMENT - VIP-100

The concern that came closest to meeting the criteria established was Threshold Technology, Inc. of Cinnaminson, New Jersey. When surveyed, they were marketing the VIP-100 an adaptive, isolated word recognition unit. One other concern, Scope Electronics of Reston, VA was developing a unit similar to the VIP-100. The VIP-100 unit was designed to automatically recognize a maximum of 64 spoken isolated word (or short continuous phrases less than 2.0 seconds duration) which could be used for data input and control applications.

The VIP-100 is comprised of four basic units. They are the Pre-Processor, the Output Display Unit, the Central Processor (Nova Mini Computer) and a model 33 ASR Teletype (Fig. 1). Speech input from the microphone or tape input is accepted by the Pre-Processor which converts it to logic signals which are then processed by the Central Processor (CP). Speech level is volume adjustable and monitored by the VIP's speech level meter. A properly positioned, special noise-cancelling microphone allows the equipment to perform accurately in both quiet and noise filled environments. The CP compares the input signal with stored references to determine which, if any, of a set of vocabulary words or short phrases were spoken. Interaction is accomplished through the Output Display Unit (ODU). If a correlation is found, an appropriate message is displayed on the ODU; a REJECT indicator illuminates if no correlation is found. Upon recognition of an input signal, a READY indicator illuminates to inform the operator that it is awaiting a new input. The model 33 ASR TTY is used for control and output/input functions. Operation of voice recognition equipment is simplified by the operating software supplied. Before using the equipment in the speech recognition mode, the equipment is first optimized to the operator's speech characteristics and for a specified set of vocabulary by use of a training

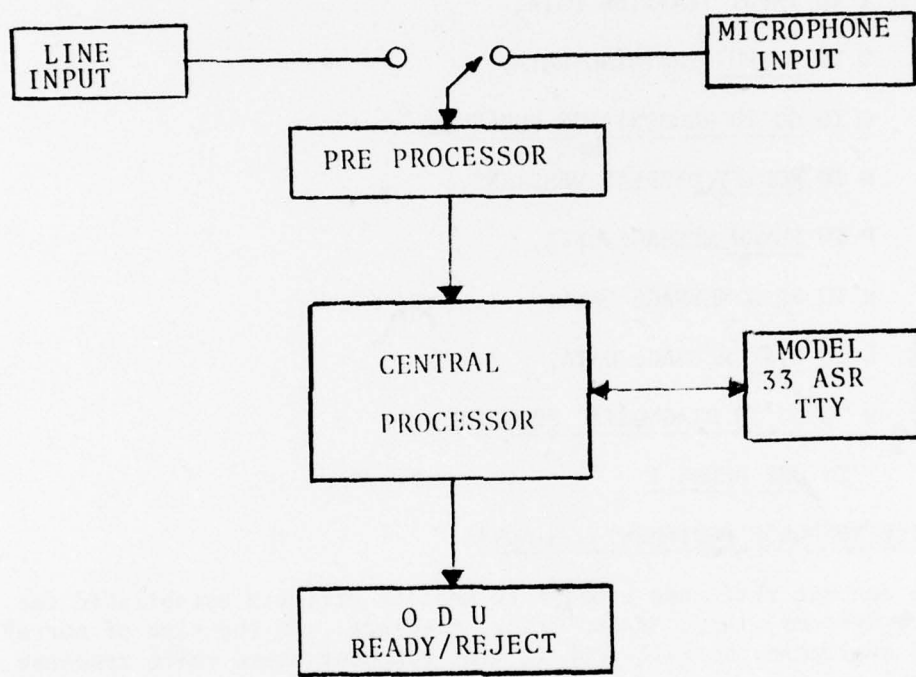


FIG. 1 VOICE RECOGNITION EQUIPMENT (VRE)

mode. To facilitate recognition, the word or phrase must have been spoken several times in the training mode while the program analyzes the pre-processed input and sets up a pattern recognition code which is used during the recognition mode (Fig. 2). The operating software includes all specified modes of operation. These modes are entered as follows:

VIP-100 WORD RECOGNITION PROGRAM

TYPE:

T TO TRAIN,

I TO INPUT TRAINING DATA,

O TO OUTPUT TRAINING DATA,

G TO GO TO RECOGNITION PHRASE,

M TO MODIFY DISPLAY MESSAGES,

P TO PUNCH MESSAGE DATA,

R TO READ MESSAGE DATA,

L TO LIST MESSAGE DATA,

D TO GO TO DIAGNOSTIC PROGRAM,

? TO USE DEBUG 1

3.3 VOICE RESPONSE EQUIPMENT - S-NOVA

The concern that came closest to meeting criteria established was *Interface Systems, Inc.*, of Ann Arbor, Michigan. At the time of survey they had available the S-11 (PDP-11 mini-computer based voice response equipment) and were nearing completion of the S-NOVA (Nova mini Computer based voice response equipment). Other equipment available from different industrial firms involved prerecorded human speech or digital waveform coding schemes with their associated disadvantages, initial cost and a requirement for large amounts of storage.

The S-Nova response equipment is a combined hardware and software package which permits phoneme programming of the CP to provide outputs in the form of English speech. Four basic units comprise the S-NOVA; they are the speech synthesizer, the audio output unit, the CP and the Model 33 ASR Teletype (Fig. 3).

Included in the speech synthesizer are an output/input interface, a synthesizer control network for output data buffering/timing control and a speech generator which converts data words electronically into analog form for conversion into the continuous sounds of speech.

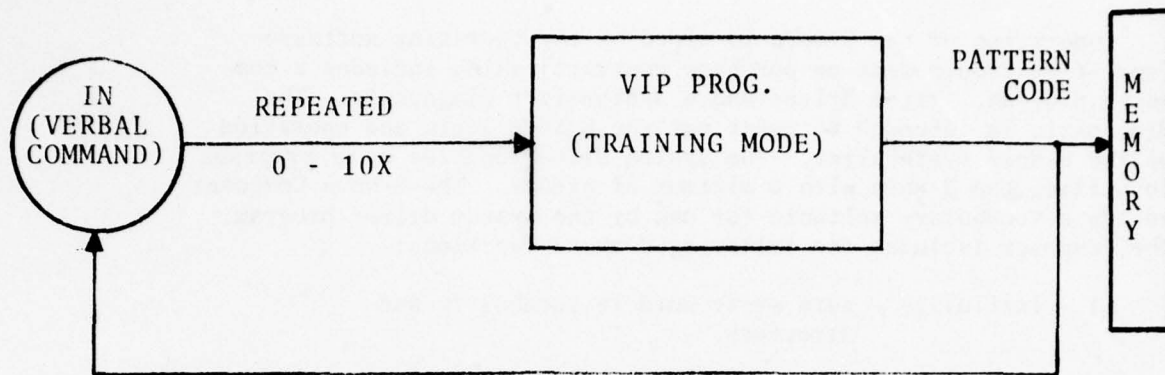


FIG. 2 VOICE RECOGNITION WORD PATTERN GENERATION

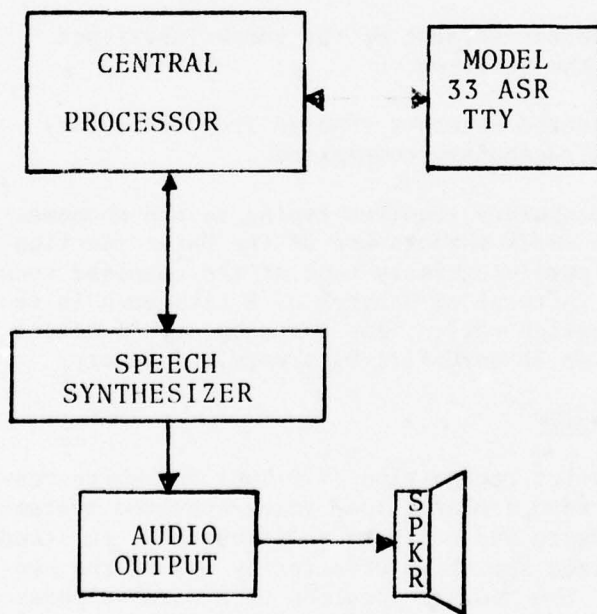


FIG. 3 SYNTHESIZED VOICE RESPONSE EQUIPMENT

Operation of the S-Nova is aided by the operating software (specified requirement on purchase contract) which includes a composer program, system driver and a Synthesizer diagnostic. The diagnostic is intended to check out the S-Nova logic and operation of the S-Nova synthesizer. The system driver enables user programs to utilize the S-Nova with a minimum of effort. The S-Nova Composer builds a vocabulary suitable for use by the system driver program. The composer includes the following Command functions:

- I - Initialize - zero every word in vocabulary and directory.
- R - Read - read vocabulary and directory from absolute binary tape.
- P - Punch - punch vocabulary and directory onto an absolute binary tape.
- E - Enter - type in the phoneme for a word or one or more phrases.
- S - Speak - selected phrases outputted through audio amplifier unit.
- L - List - selected segment of the vocabulary typed at the teletype.
- D - Delete - selected segments removed from vocabulary and vocabulary compressed.

Composing a given vocabulary requires typing in the phonemes that make up intelligible words through use of the Enter function (Fig. 4). On Command an absolute binary tape of the composed vocabulary can be generated. A total of 9 bytes of 8 bits each is required for the average English word. Thus a vocabulary of almost 1800 words can be stored in 8K words (16 bits/word) of memory.

3.4 CLOSED LOOP VOICE SYSTEM

Integration of the voice recognition (VIP-100) and voice response (S-Nova) equipment into a closed loop voice operated system was accomplished via hardware and software modification. In stand alone operation, synthesized speech is effected by use of the synthesizer driver program. The storage location of a word or phrase to be spoken is accessed by typing in a decimal (0 to ∞) number pseudo. Phrase or word phoneme data stored in the defined location is taken from memory by the synthesizer driver program and used to drive the

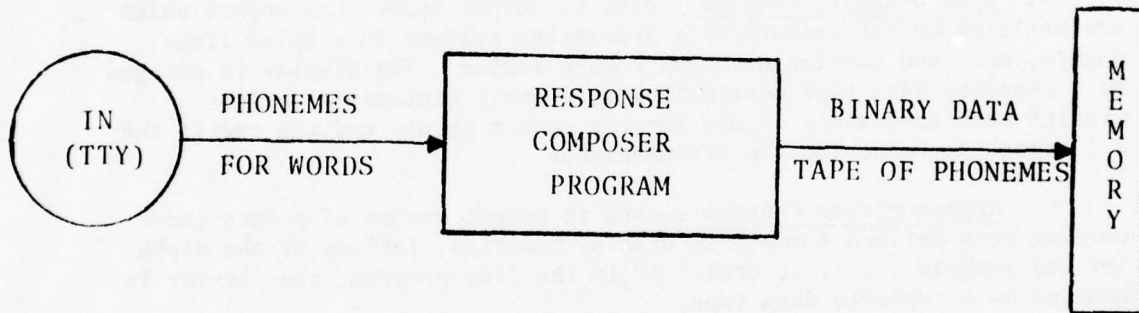


FIG. 4 VOICE RESPONSE SYNTHESIZER DATA GENERATION

synthesizer device. Modification of the VIP program involved taking the recognized word or phrase number (integer 0 to 63) generated and inserting it into the synthesized Data Organization program. The synthesized Data Organization program converts the integer to the decimal pseudo for the appropriate response thereby eliminating the typing requirement. Closed loop operation was completed by modifying the synthesizer device program to yield control to the VIP program upon completion of an utterance by the speech generator (Fig. 5).

3.5 ASSOCIATED PERIPHERALS

To enhance the man-machine interaction several peripherals were incorporated into the system.

Full graphic/alphanumeric visual capability was implemented through use of a vector/point generating system and a cathode ray tube (CRT) oscilloscope type monitor. Efficient compatibility with the Voice system was realized by developing the following assembly language programs which were incorporated as subroutines.

a. Line Graphics Program - used to output individual points which are utilized by the vector/point generating systems to display lines, points, arcs and circles with any chosen radius. The display is changed by a separate data tape designating x, y and z (intensity) values. Circles require x and y values for the center points and the radii; the z (intensity) value for the circumference.

b. Alphanumerics Program - used to output groups of points (according to a defined format) to display numerics, letters of the Alphabet and symbols (+, -, →, etc). As in the line program, the display is changed by a separate data tape.

c. Arrow (Blinking Program) - used to blink the arrow symbol to attract the operators attention to a specific part of the display. May easily be modified to blink any symbol.

d. Erase Program - used to clear monitor of display on Voice command.

In addition to the full graphic visual capability, supplementary alphanumeric capability was added through adaptation of a character dot matrix generating system and a high resolution television (TV)-type monitor. The software facilitates message display via a set of subroutines which permit straight-forward control of text display (Supplementary Display Program).

Advantages offered by the addition of visual display capability include more efficient, effective and reliable man-machine interaction. Visual feedback reinforces the audio feedback (speech response) through

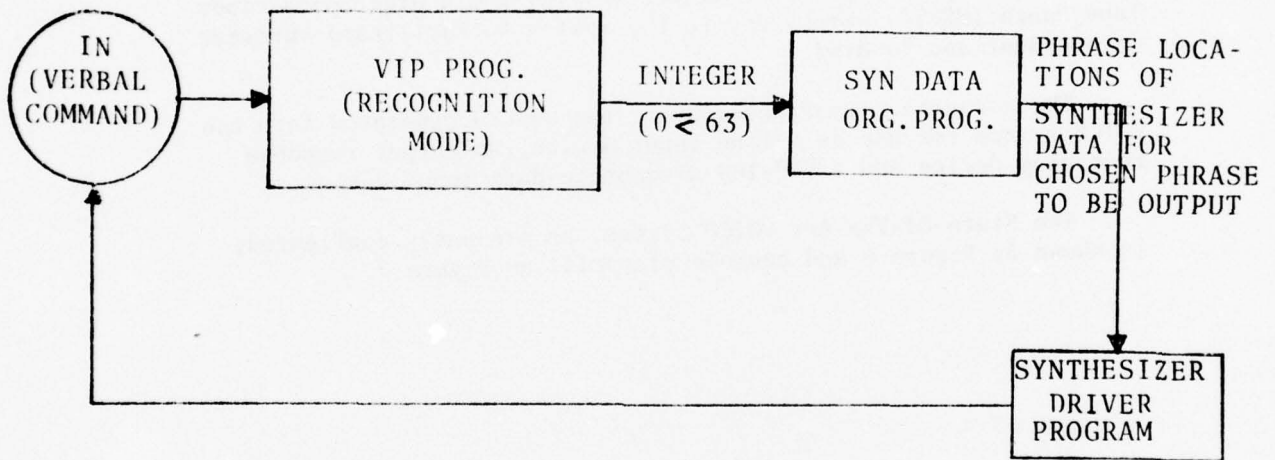


FIG. 5 CLOSED LOOP VOICE O/I PROGRAM

the use of pictures, diagrams, schematics and operational procedures. Disadvantages include restriction of operator mobility to display line of sight area and higher system cost due to additional peripheral expenses.

Hard copy capability was added by interfacing an analog plotter to the system via the graphic/alphanumeric display generating system. The high-speed output of the display interface necessitated development of a software routine that adjusts the output speed to conform to speeds attainable by the analog plotter.

Printed copy capability was added through utilization of the model 33ASR's printer mode. Hard copy provides an efficient and very reliable means of recording pertinent information. A disadvantage is the time required to generate the hard copy. This can be easily overcome with more elaborate recording equipment.

A High Speed Paper Tape Reader (HSPTR) and a High Speed Paper Tape Punch (HSPTP) were added to the system to facilitate software development and loading.

The cassette player/recorder system was incorporated into the VOICE system for use as a line input device, an output response recording device and a VIP-100 diagnostic data input device.

The State-Of-The-Art VOICE System, as presently configured, is shown as Figure 6 and console pictorial as Figure 7.

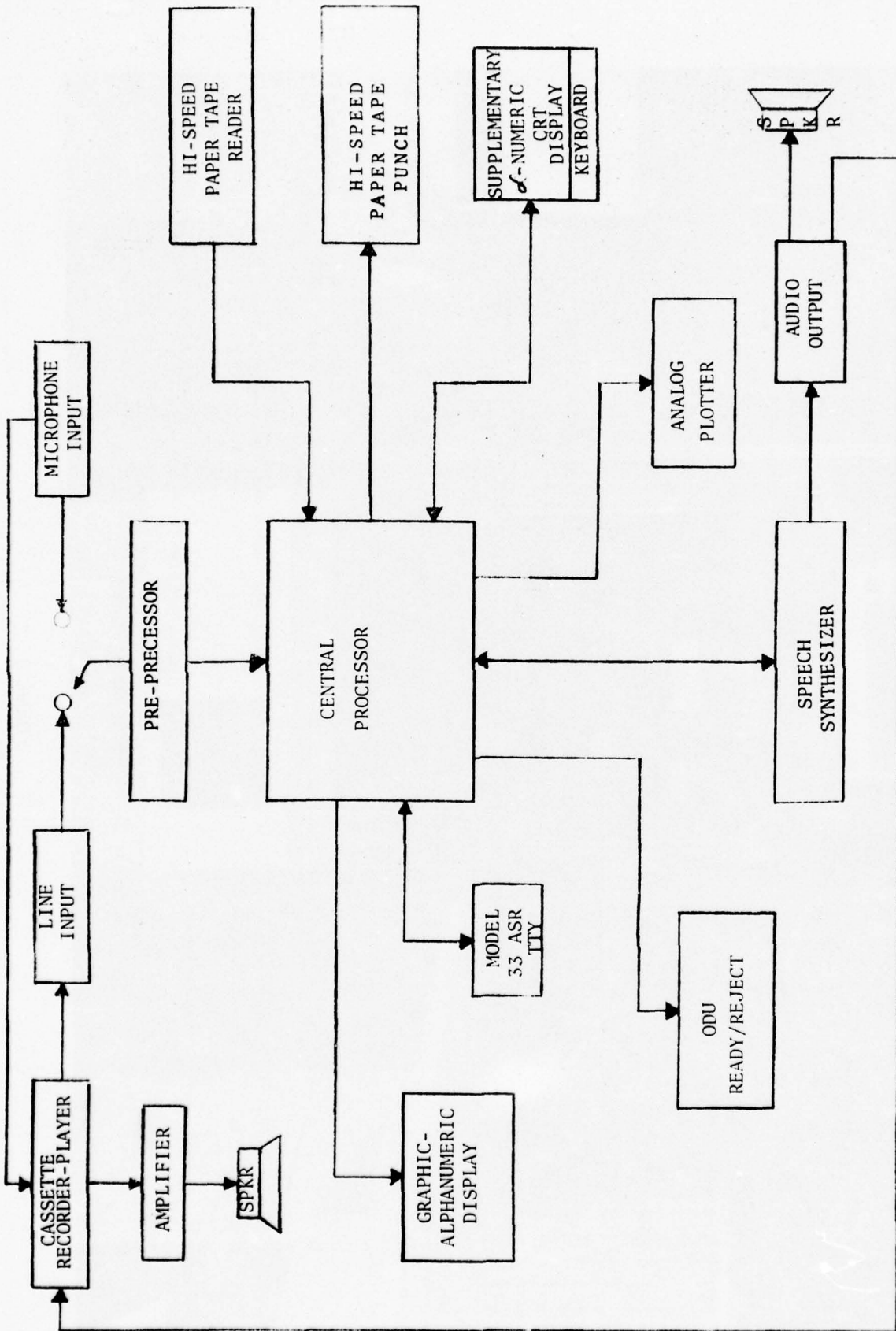


FIG. 6 - STATE-OF-THE-ART VOICE SYSTEM

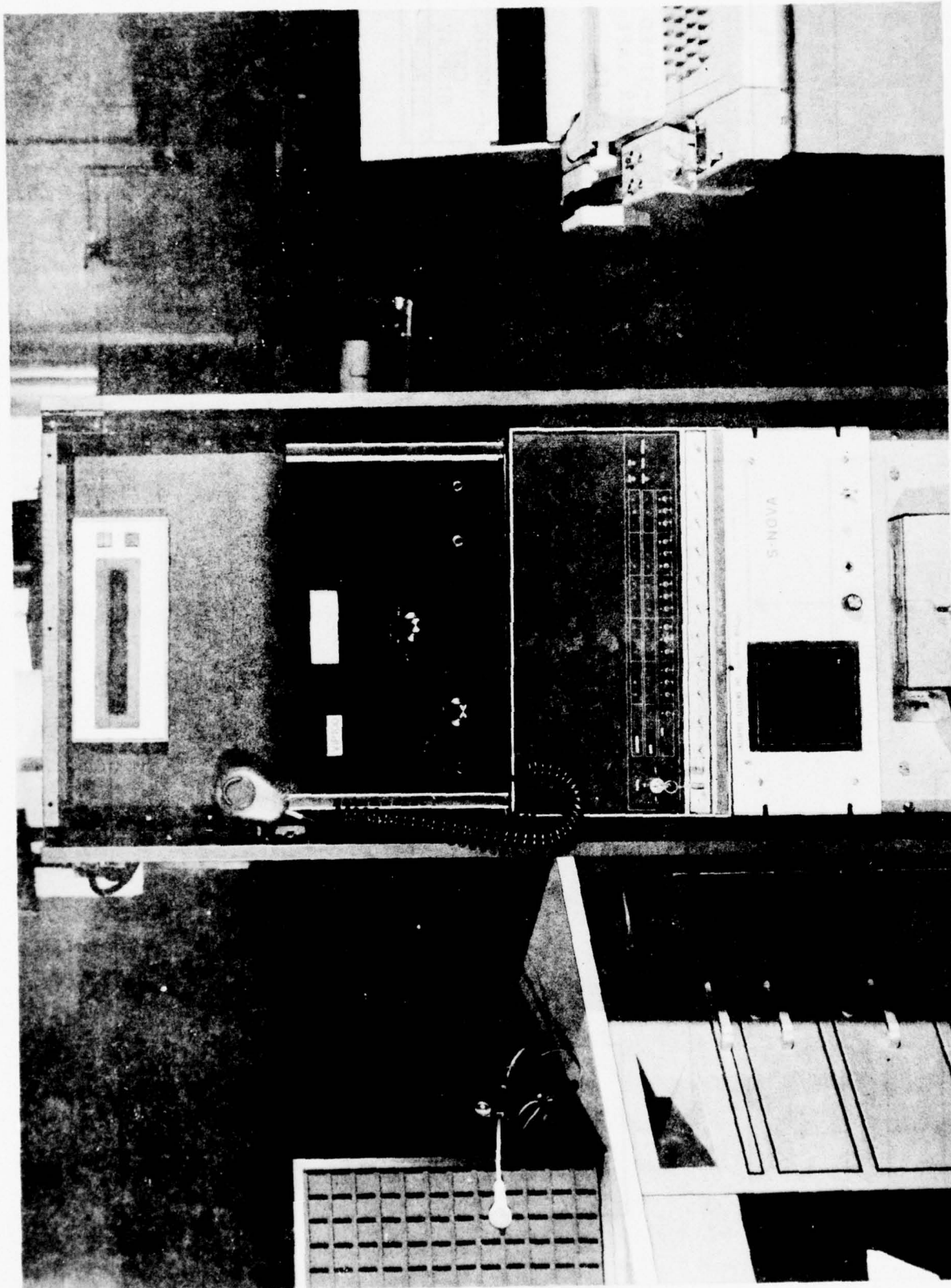


FIGURE 7

SECTION IV

VOICE SYSTEM EVALUATION

4.1 SYSTEM EVALUATION AND DEVELOPMENT

a. Voice Recognition

The voice recognition equipment is an adaptive unit requiring prior optimization for both the words or phrases selected and the operator's particular speech characteristics through use of a training routine.

Recognition accuracy is influenced by many factors. Any temporary changes in the operator speech characteristics which may result from colds, respiratory ailments, placement of microphone, increase or decrease in tone of voice, and pronunciation may, in some instances, be detrimental to recognition accuracy. Because basic speech components are analyzed, abnormalities due to colds, sore throats, or hoarseness normally will not hamper recognition of the operator's speech. On occasions when these effects are detrimental, the problem is easily and quickly resolved by retraining. Operation in a noise-filled environment is accomplished through use of a close talking noise-cancelling microphone worn suspended from a lightweight headband. Positioning and alignment of the head mounted microphone proved critical and too inconvenient for general use so a lightweight, hand held noise-cancelling microphone with lip piece was substituted. Overall recognition accuracies were not affected. While the noise-cancelling microphone is the optimum compromise between reducing background noise and obtaining high quality speech, it introduces the factor of extraneous signals caused by breath noise. A strong tendency to exhale at the end of voiced inputs produces signal levels in the microphone comparable to speech levels. The problem was alleviated by modifying a circuit in the pre-processor.

Full rejection of breath noise can be effected by proper coordination of the push-to-talk switch (an integral part of the hand held microphone).

The limiter amplifier, which reduces gain for the strongest speech inputs, was bypassed since it was determined that overall recognition accuracy was diminished. Maximum accuracy is realized by distinct pronunciation with the volume control adjusted for an average level five on the speech level meter (scaled 0 to 10) while speaking in a normal tone during both the training and recognition phases.

Experiments using the voice recognition unit as a "universal" recognition unit, were conducted with encouraging recognition accuracies recorded. During the training phase each of five speakers trained the unit for a given word or phrase twice. The features extracted were automatically averaged, classified, and stored. When operating in the recognition mode, the same words spoken by a wide range of different

TABLE I - PHONEME LIST

| <u>PHONEME</u> | <u>EXAMPLE</u> | <u>PHONEME</u> | <u>EXAMPLE</u> |
|----------------|------------------|----------------|------------------|
| AW | <u>L</u> AW | UH | BU <u>T</u> |
| IE | ZE <u>R</u> O | AE | CA <u>T</u> |
| IE, E | TH <u>RE</u> E | W | W <u>O</u> N |
| EH | TE <u>N</u> | R | TH <u>RE</u> E |
| P | <u>P</u> ENNY | Z | <u>Z</u> ERO |
| AH | CA <u>R</u> | SH | <u>S</u> HOW |
| YI | <u>Y</u> ES | T, CH | <u>CH</u> AIR |
| A, AY | <u>E</u> IGHT | V | SE <u>V</u> EN |
| S | <u>S</u> IX | B | <u>B</u> ED |
| D | <u>D</u> O | N | <u>N</u> INE |
| F | <u>F</u> OR | M | <u>M</u> ILE |
| ER | WEA <u>TH</u> ER | T | <u>T</u> EN |
| TH | <u>TH</u> REE | Y | S <u>IX</u> TY |
| THV | <u>TH</u> EN | I | <u>S</u> IX |
| IU, U | <u>T</u> WO | O | <u>N</u> O |
| G | <u>G</u> ET | AH, IY | <u>N</u> INE |
| H | <u>H</u> ELLO | EH | <u>E</u> D |
| D, J | <u>E</u> DGE | ZH | A <u>Z</u> URE |
| K | <u>C</u> AME | IH | STA <u>T</u> ION |
| L | <u>H</u> ELLO | I, IH | <u>S</u> IN |
| OO | <u>B</u> OOK | OO, OOH | <u>B</u> USH |
| NG | <u>R</u> ING | Q | <u>Q</u> ICK |

The user quickly discovers that word composition is somewhat subjective and, therefore, one representation will sound better to one person than another. Since the choice of inflections is often related to context, the user should attempt to phonetically represent the text desired as he would deliberately pronounce it. Normal stress will apply in most instances. At times more or less is desired. When greater stresses than those produced by prefixing and suffixing phonemes are required, the user may repeat the phoneme. Accenting is implemented by a word separation pause. Familiarity with the sounds of the individual phonemes and the combination required to produce the many glides which are so common in the English language is indispensable to efficient speech composition (Table 1).

The S-Nova was modified to correct occasional erratic speech pronouncements and random slowing of the speech output. Intelligibility of the speech response was improved by increasing the speech output rate. Throughout the process of generating speech response, the criterion for selection of the phoneme representation has been intelligibility rather than aesthetics.

4.2 VOICE FEASIBILITY DEMONSTRATIONS

Potential applications capability of VOICE was demonstrated by four exercises. Demonstration one illustrates man-machine interface as applied to circuit board testing. An audio amplifier with easily changed components was selected as the circuit board to be tested. Performance was analyzed through use of A/D conversion and digital encoding. The amplifier is inserted into the system and analysis is initiated by the operator who commands the computer to commence testing. The operator is then instructed to complete a preliminary set up. The graphics/alphanumeric display is utilized to generate a layout pictorial of the amplifier. Specific components are further identified, by blinking arrow pointers, as required. External equipment connections are also identified by blinking arrow pointers. The supplementary alphanumeric display is used for listing parts, voltage requirements, amplifier input/output parameters, amplifier conditions, faulty components, and other pertinent information. Voice response instructs the operator in completing the preliminary set-up, diagnosing problems, and taking corrective action. The procedure with man-machine interaction is described in Table II - Audio Amplifier Circuit Test. Figure 8 illustrates a hard copy (analog plotter) of the audio amplifier circuit displayed on the graphic/alphanumeric display.

Demonstration two illustrates man-machine interface as applied to acoustical diagnostics. While the voice response equipment was designed to process basic speech components, the evaluation process indicated a limited capability of processing acoustical signals within the designed bandwidth. A unique potential application of the system could be acoustical fault diagnostics. The feasibility of such an application is demonstrated by an internal combustion engine tune-up. Since a competent experienced mechanic can diagnose various engine malfunctions by the acoustics emitted, it naturally follows that a properly trained machine could do likewise. The source of acoustics chosen was a readily available M151 Engine.

TABLE II - AUDIO AMPLIFIER CIRCUIT TEST

| VIP-100 INPUT | OUTPUT DISPLAY UNIT (ODU) | GRAPHIC ALPHA - NUMERIC DISPLAY | SUPPLEMENTARY ALPHA-NUMERIC DISPLAY | S - MCVIA OUTPUT SPEECH RESPONSE |
|--|---------------------------|--|--|--|
| Computer, perform Amplifier Test | YES, SIR | | | YES, SIR. Which amplifier do you wish to test? |
| Test M1 Audio Amplifier | OKAY | Picture of Audio Amplifier (see Figure 8) | M1 Audio Amplifier Parts List. Power Requirements. Amplifier I/O parameters. | OKAY. Preliminary set-up is as follows: Connect channel 1 input probe of oscilloscope from A to B on circuit board. Connect channel 2 input probe of oscilloscope from C to D on circuit board. What kind of waveform do you wish to be used as a stimulus? |
| Use a 200 millivolt Sine Wave | 200 MV | Blink arrow pointers at A & B | | What frequency? |
| One KiloHertz | 1 KHz | Blink arrow pointers at C & D | | Preliminary Set-up completed. |
| Analyze circuit | ACCEPTED | All arrow pointers erased. Picture of Audio Amplifier without pointers. | All previous data plus Audio Amplifier accepted. | Audio Amplifier meets acceptance criteria. Amplifier Test completed. |
| IF THE COMMAND ANALYZE CIRCUIT GIVES THE ABOVE RESULTS, AUDIO AMPLIFIER IS WITHIN SPECIFICATIONS. | | | | |
| ERASE | ODU CLEARED | Picture remains | All data to this point remains | READY |
| Analyze circuit | REJECT | Picture with pointer to faulty component | All previous data plus faulty component | Problem Identification and correction. |
| IF THE COMMAND ANALYZE CIRCUIT GIVES THE ABOVE RESULTS, AUDIO AMPLIFIER IS FAULTY. IF THE PROBLEM IS C1, INTERACTION IS AS FOLLOWS: | | | | |
| | | Blinking arrow pointer to C1 | C1 open. Replace with 0.1 ufd, 25 wvdc capacitor. | Audio amplifier has parasitic oscillations. Decoupling capacitor C1 open. Turn off power and replace with a 0.1 ufd capacitor. Turn on power. |
| Analyze circuit | ACCEPTED | Picture of Audio Amplifier without pointers | All previous data plus Audio Amplifier accepted | Audio Amplifier meets acceptance criteria. Amplifier Test completed. |
| IF THE PROBLEM IS IC1, INTERACTION IS AS FOLLOWS: | | | | |
| | | Blinking arrow pointer to IC1 | IC1 has internal short circuit. Replace with LM 380 | Audio Amplifier has no output. Analysis indicates internal short. Turn off power and replace IC1 with LM380. Turn on power. |
| Analyze circuit | ACCEPTED | Picture of Audio Amplifier without pointers | All previous data plus Audio Amplifier accepted. | Audio Amplifier meets acceptance criteria. Amplifier Test completed. |
| IF THE PROBLEM IS R3 INTERACTION IS AS FOLLOWS: | | | | |
| | | Blinking arrow pointer to R3 | R3 open. Replace with 2.7 ohm, 1/2 watt resistor | Audio Amplifier has a high frequency, small amplitude oscillation of approximately 10 mega Hertz during its negative swing. This oscillation, while not detrimental to amplifier quality, should be suppressed when operating in an environment sensitive to RF. Turn off power and replace R3 with 2.7 ohm, 1/2 watt resistor. Turn on power. |
| Analyze circuit | ACCEPTED | Picture of Audio Amplifier without pointers | All previous data plus Audio Amplifier accepted | Audio Amplifier meets acceptance criteria. Amplifier Test completed. |
| ERASE DISPLAYS | --- | --- | --- | READY |

AUDIO AMPLIFIER

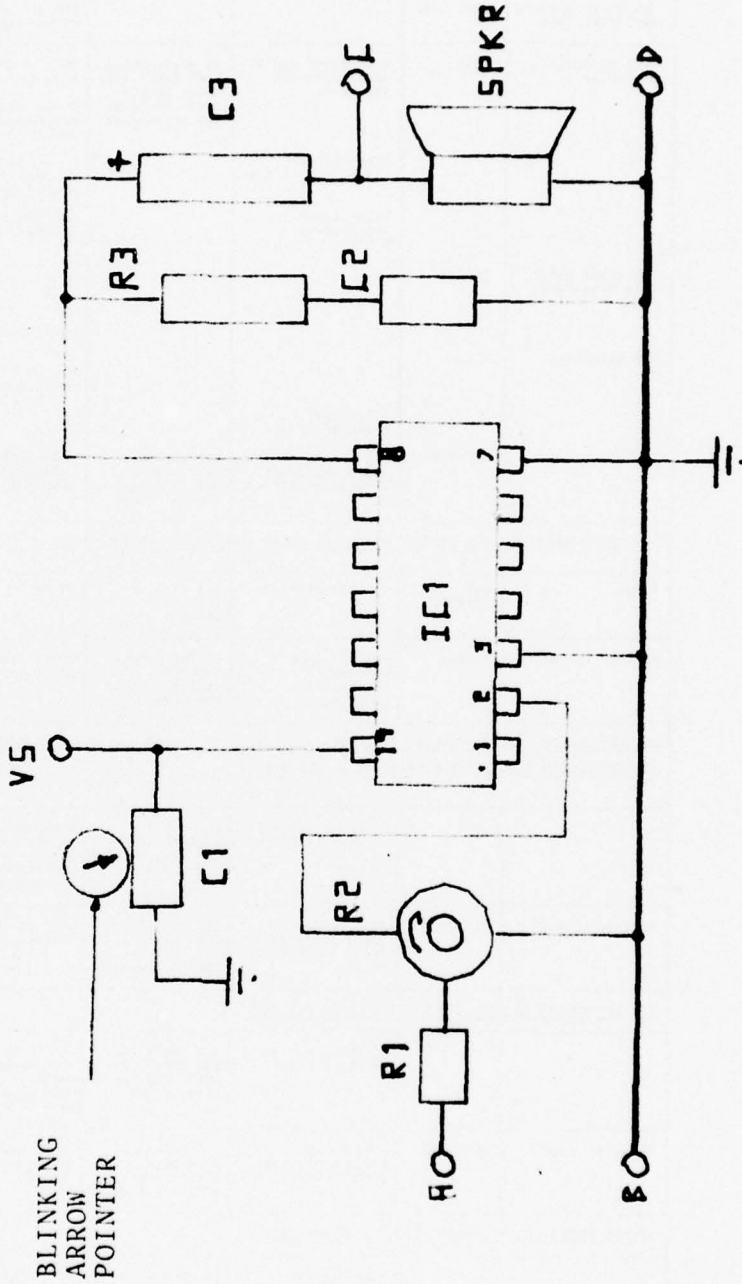


FIG. 8 - AUDIO AMPLIFIER CIRCUIT

The engine was adjusted for various malfunctions which would emit acoustical signatures that would be acceptable to the VIP-100 (e.g., spark plug misfire, idle too high, idle mixture too rich, and timing advanced). After the system had been trained to respond to acoustics produced by the malfunctions, an engine tune-up diagnostic procedure was performed. The operator initiated the procedure by instructing the computer to perform the M151 Engine analysis. Ensuing interaction resulted in detection and correction of the engine problems in proper sequence. The graphic alphanumeric display was used to illustrate the engine, faulty components, adjustment screws, and other information pertinent to the diagnostic procedure.

Upon detection of engine problems, the operator was given the choice of correcting the problems or instructing the computer to print out a list of the problems with recommended corrective action. If the operator chooses to take the corrective action himself, he requests instructions from the system and is guided through a step-by-step remedial process which results in a tuned engine. The procedure is described in Table III - Engine Diagnostic Procedure. Figure 9 illustrates a hard copy (analog plotter) of the M151 Engine layout displayed on the graphic/alphanumeric display. The supplementary display could be used for listing the procedure and giving further detailed instructions.

Demonstration three illustrates man-machine interface as applied to fire mission control by a forward observer. Once the system has been trained to respond to a given forward observer, it rejects inputs from anyone else, thereby, assuring security of the system. Forward observer fire direction center interactive communication can be accomplished by use of a wireless transceiver reinforced by a miniature head mounted display. The same kind of communication loop can be established between the battery commander and the fire direction center. A sample fire mission dialogue is described in Table IV.

Demonstration four illustrates use of the VOICE System as a multi-function scientific calculator. Instructions and data are spoken into the system which performs a chain calculation and outputs the answer via the ODU and the speech synthesizer. An entry error is cleared by the spoken command CLEAR ENTRY. The spoken command CLEAR zeros all the registers. Operation of the calculator is best described by use of an example. Given the following problem:

$$X = \frac{(A+B-C)D}{E} - F + \sqrt{G}$$

$$A = 4 \quad C = 2 \quad E = 5 \quad G = 2.5$$

$$B = 3 \quad D = 15 \quad F = 30$$

The procedure is described in Table V.

TABLE III M151 ENGINE DIAGNOSTIC PROCEDURE

| VIP-100 INPUT | OUTPUT DISPLAY UNIT (ODU) | GRAPHIC ALPHA-NUMERIC DISPLAY | PRINTER LISTING | S-NOVA OUTPUT SPEECH RESPONSE |
|--|---------------------------|---|--|--|
| Computer perform engine test | YES SIR | | | Yes Sir. Which engine do you wish to test? |
| Test M151 Engine | OKAY | M151 Engine Test | | Okay. Let me listen to the engine idle |
| Engine idle with malfunctions | REJECT | M151 Engine Test Engine has multiple problems as follows: 1. Spark Plug #1 misfire 2. Timing advanced. 3. Idle mixture too rich. 4. Idle too high. | | Engine has multiple problems as follows: 1. Spark Plug #1 misfire. 2. Timing advanced. 3. Idle mixture too rich. 4. Idle too high. Do you wish to take corrective action? |
| No | LISTING | As above | M151 Engine has multiple problems as follows: 1. Spark Plug #1 misfire. Replace with new spark plug with gap set to .030. 2. Timing advanced. Set timing to 6° before top dead center. 3. Idle mixture too rich. Set mixture adjustments for proper Air-Fuel ratio. 4. Idle too high. Set idle adjustment for 600 RPM. | |
| YES. Instruct me. | REJECT PLUG MISS | M151 Engine Test Illustration of M151 engine with blinking arrow pointing to Spark Plug #1 (see Fig. 9). | | Corrective action will be accomplished in the proper sequence. Replace Spark Plug #1 with new spark plug with gap set to 30 thousandths. Refer to pointer on display. Let me listen to the engine again. |
| Engine idle with remaining malfunctions. | REJECT TIMING ADV | As above with blinking arrows pointing to timing mark and distributor. | | Adjust timing. Refer to pointers. Set timing to 6° before Top Dead Center. Let me listen to the engine again. |
| Engine idle with remaining malfunctions. | REJECT RICH-HIGH | As above with blinking arrows pointing to idle mixture adjustment and idle adjustment | | Adjust idle and mixture. Refer to pointers. Set mixture adjustments for proper air - fuel ratio. Set idle adjustment for 600 RPM. Let me listen to the engine again. |
| Engine idle (In Tune) | ACCEPTED | As in Figure 9 with no blinking arrows. | | Engine problems corrected. Engine meets acceptance criteria. Engine Test completed. |
| ERASE DISPLAY | --- | --- | --- | READY |

M151 ENGINE TEST

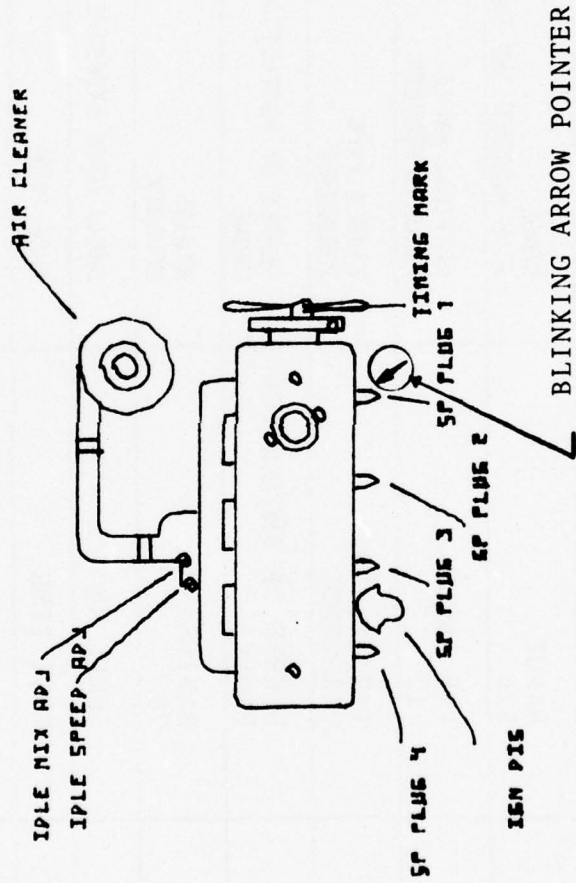


FIG. 9 - M151 ENGINE ILLUSTRATION

TABLE IV FIRE MISSION - FORWARD OBSERVER/FIRE DIRECTION CENTER

| VIP-100 INPUT FORWARD OBSERVER | OUTPUT DISPLAY UNIT (ODU) | S-NOVA OUTPUT SPEECH RESPONSE |
|--------------------------------|----------------------------|---|
| FIRE MISSION | FIRE MISSION | FIRE MISSION |
| DIRECTION THREE TWO ZERO ZERO | DIR 3200 | DIRECTION THIRTY TWO HUNDRED |
| RANGE FIVE TWO ZERO | RANGE 520 | RANGE FIVE HUNDRED AND TWENTY |
| VERTICAL ANGLE MINUS ONE SEVEN | VER ANGLE -17 | VERTICAL ANGLE MINUS SEVENTEEN |
| TARGET TYPE PERSONNEL | TARGET TYPE PERSONNEL | TARGET TYPE PERSONNEL |
| DEGREE OF PROTECTION PRONE | DEGREE OF PROTECTION PRONE | DEGREE OF PROTECTION PRONE |
| RADIUS SEVEN ZERO | RADIUS 70 | RADIUS SEVENTY |
| SHELL HE | SHELL HE | SHELL HIGH EXPOSIVE |
| FUSE TIME | FUSE TIME | FUSE TIME |
| CONTROL | CONTROL | CONTROL |
| WHEN READY | WHEN READY | WHEN READY |
| COMPUTE | COMPUTE | * DEFLECTION-THREE FOUR ONE NINER (3419) TIME - ONE FOUR POINT SEVEN (14.7) QUADRANT - TWO TWO FOUR (224) |
| TARGET DESTROYED | CEASE FIRE E.O.M. | * CEASE FIRE END OF MISSION |

* SENT TO BATTERY

TABLE V - SAMPLE CALCULATION

| VIP-100 SPEECH INPUT | COMPUTER, PERFORM CALCULATION | OUTPUT DISPLAY UNIT (ODU) | S-NOVA SPEECH RESPONSE |
|-------------------------------|-------------------------------|------------------------------|--|
| COMPUTER, PERFORM CALCULATION | | YES SIR | YES SIR. INSTRUCT ME |
| FOUR | 4 | | FOUR |
| PLUS | 4+ | | PLUS |
| THREE | 7 | | THREE |
| MINUS | 7- | | MINUS |
| TWO | 5 | | TWO |
| TIMES | 5 X | | TIMES |
| ONE, FIVE | 75 | | ONE, FIVE |
| DIVIDED BY | 75 ÷ | | DIVIDED BY |
| FIVE | 15 | | FIVE |
| MINUS | 15- | | MINUS |
| THREE, ZERO | -15 | | THREE ZERO |
| PLUS | -15+ | | PLUS |
| SQUARE ROOT OF | -15+ $\sqrt{\quad}$ | | SQUARE ROOT OF |
| TWO, POINT, FIVE | -13.418861 | | TWO, POINT, FIVE |
| EQUALS | -1.3418861X10 ⁰¹ | | NEGATIVE ONE, POINT, THREE, FOUR ONE, EIGHT, EIGHT, SIX, ONE, TIMES, TEN, TO THE ZERO, ONE |

4.3 GENERAL PROGRAM FLOW OF DEMONSTRATIONS

Verbal and acoustical inputs are preprocessed by the VIP-100 and the results analyzed by the VIP-100 Program. The integer derived from the input data is then inserted into the select program which controls the sequence of peripheral selection for the proper cycle (Figure 10). Display generating programs are selected as required. Following the display requirements, the S-Nova Synthesizer Program takes over and outputs the appropriate message. Upon completion of the speech response, control is returned to the VIP-100 Program.

4.4 NEW DEVELOPMENTS IN VOICE OUTPUT/INPUT

a. Voice Recognition

Threshold Technology, Incorporated has developed more compact VOICE recognition equipment based on the Digital Equipment Corporation LSI-II Processor. The LSI-II offers minicomputer performance in a microcomputer package and price. Miniaturization of the processor circuitry was realized through use of integrated circuits wherever possible. Threshold has developed software that further classifies distinguishing phonemes, thereby eliminating recognition problems involved with using vocabulary sets containing one syllable words that rhyme. Hardware has been modified to preclude acceptance of breath noise. In addition to the material handling and quality assurance inspection applications, new dimensions have been added through VOICE programming for numerical control.

Traditionally, human interface problems associated with programming and software have been a major bottleneck in computer-based numerical control systems. Voice programming for numerical control (VNC) allows factory personnel, using normal English words, to speak commands needed for parts programming. All the operator requires is a knowledge of blueprints and basic machine tool operations. Threshold is researching the formidable task of recognizing continuous speech. Future plans call for an Intel 8080 Microprocessor Based Voice Recognition System which should open up many new applications areas.

Threshold Technology, Incorporated has recently completed work on an advanced development model of a Voice Input Code Identifier (VICI) for the Rome Air Development Center. The VICI is an isolated word recognition unit based upon the VIP-100. Tailored hardware and software modifications were made to the VIP-100 to accomplish recognition of the VICI vocabulary (English digits 0 - 9 and four control words, CANCEL, ERASE, VERIFY, and TERMINATE) for a large male speaker population without adaptation or training for any given speaker. Individual recognition accuracy rates of 98% were realized. Incorporation of an output display unit enables the speaker to verify that voiced inputs have been correctly identified.

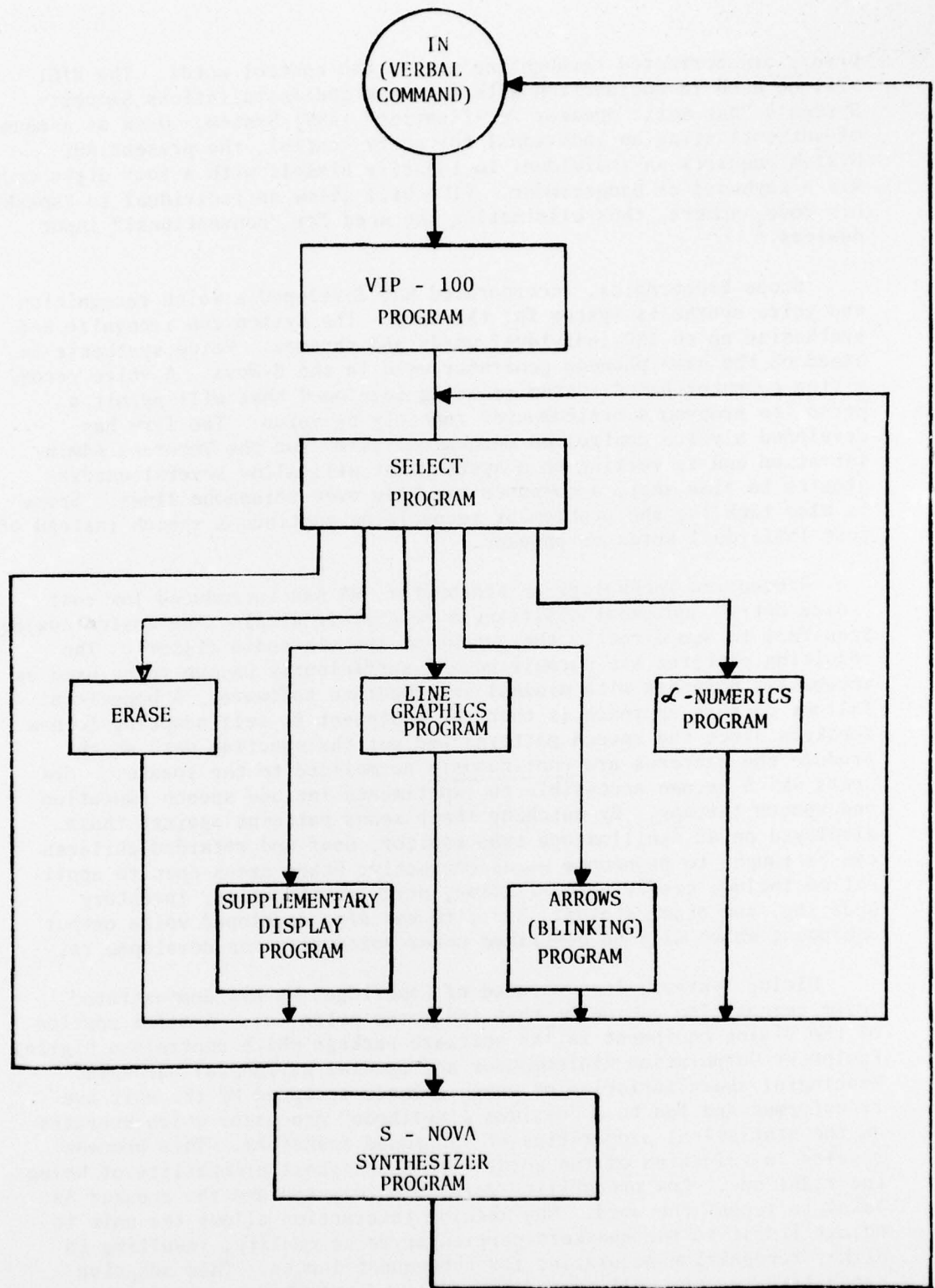


FIG. 10 - GENERAL PROGRAM FLOW OF DEMONSTRATIONS

Errors are corrected through the use of the control words. The VICI will be used in conjunction with the Base and Installations Security System's "Automatic Speaker Verification" (ASV) System. Used as a means of authenticating an individual for entry control, the present ASV System requires an individual to identify himself with a four digit code via a keyboard or badge reader. VICI will allow an individual to "speak" his code numbers, thus eliminating the need for "conventional" input devices.³

Scope Electronics, Incorporated has developed a voice recognition and voice synthesis system for the Navy. The system can recognize and synthesize up to 150 individual words and phrases. Voice synthesis is based on the same phoneme generator used in the S-Nova. A voice recognition computer-based system is being developed that will permit a person to program a minicomputer remotely by voice. The firm has developed a voice controlled mechanical "arm" for the Veterans Administration and is working on a system that will allow several quadriplegics to time share a computer remotely over telephone lines. Scope is also tackling the problem of recognizing continuous speech instead of just individual words or phrases.

Preception Technology of Winchester, MA has introduced low cost "voice entry" equipment utilizing a relatively simple sine-cosine analog transform to sum directly the output of its six audio filters. The resulting patterns are normalized and sufficiently unique to be used as vocabulary pointers with minimal intermediate software. A beneficial fallout of this approach is that the equipment is self-adapting to new speakers since the speech patterns and not the spectral data which produce the patterns are continuously normalized to the speaker. New areas which become accessible to experiments include speech education and speech therapy. By matching their sound patterns against those displayed on an oscilloscope type monitor, deaf and retarded children can be taught to pronounce words correctly. Other areas open to application include credit authorization, price verification, inventory updating, and order status. The firm has also developed voice output equipment which will be described under voice response developments.

Dialog Systems, Incorporated of Cambridge, MA has demonstrated voice recognition equipment that indicates potential. A major portion of the Dialog equipment is its software package which controls a Digital Equipment Corporation minicomputer and special peripheral hardware. Meaningful characteristics of voice signals accepted by the unit are transformed and fed to a "maximum likelihood" processor which operates on the statistical properties of the sound transform. This process results in selection of the words with the highest probability of being the right one. Low probability words are rejected and the speaker is asked to repeat the word. Man-machine interaction allows the unit to adjust itself to the speakers particular voice quality, resulting in higher recognition accuracies for subsequent inputs. This adaptive capability is very well suited to radio and telephone communication links.

³P.B. Scott, Voice Input Code Identifier. Tech Report #RADC-TR-75-188, July 1975.

Dialog also is working on the continuous speech problem. Recently Dialog has announced that it will be concentrating on applications for its new continuous speech recognition equipment. Instead of identifying whole words, the equipment identifies syllables and eliminates the need for pauses between words. The speech computer then determines which syllable combinations make up which words in a vocabulary set. The continuous speech recognition equipment, expected to be introduced in the late fall of 1976, has a recognition accuracy objective of greater than 95%. Adaptive capability will be a standard feature.

The Defense Advanced Research Projects Agency currently has a major five year program of research of the analysis of continuous speech by computer being performed by several contractors. Stanford Research Research Institute (SRI) has been participating in the program with other ARPA contractors. The long term objective of the research at SRI on speech understanding is to develop the technology that will allow speech understanding systems to be designed and implemented for a wide variety of different task domains and environmental constraints.

Currently, SRI is working cooperatively with the System Development Corporation on the design and implementation of a joint system.

Two task domains have been selected for the duration of the current five year program:

1. Data management of a file containing information about selected ships from the fleets of the United States, the Soviet Union, and the United Kingdom.

2. Maintenance of electro-mechanical equipment in a work station environment with the system as a computer consultant. A "milestone system" is scheduled for on-site demonstration in early fall 1976.

Other contractors participating in this major program of research include Carnegie-Mellon University and Berkley Campus, UCLA.

Voice recognition equipment is being evaluated in many diverse applications. Included are the following:

1. Security systems using automatic speaker identification.
2. Manufacturing, inspection, and quality control.
3. Postal service parcel routing via spoken zip code.
4. Law enforcement and military computer communication links.
5. Voice controlled motorized mechanical "arms" and wheel chairs and typewriters for the handicapped.
6. Various sorting and handling operations.

7. Computer programming through remote voice input.
8. Speech education and speech therapy.

b. Voice Response

The speech generator portion of the S-Nova has gone through an evolutionary process. Marketed by Vocal Interface Division, Federal Screw Works of Framingham, MA, the latest version (VS-6) has been optimized for a "standard" American English dialect. Programming for any word or phrase is considerably easier than in previous synthesizer models.

Proper allophones, transitions, and inflections are generated automatically in the majority of cases. Unlike previous synthesizer models, the VS-6's inflection is completely independent from phoneme timing. The VS-6 is designed so that nominal pitch falls between inflection levels 2 and 3. Use of inflection 2 or 3 will result in good intelligible speech response. Inflection can be controlled by software allowing the user to assemble sentences out of a pre-stored vocabulary set and have given words inflected according to sentence grammar. Since intelligibility is enhanced by short gaps between words, the synthesizer is designed to generate continuous speech with appropriate gaps. Speech rate and pitch controls give the unit additional flexibility. Overall, the speech response quality is considerably improved.

Master Specialties Company of Costa Mesa, CA has developed a proprietary technique for digitizing and storing whole words in Metal Oxide Semiconductor (MOS) Read Only Memories (ROM). Analog audio signals are converted into digital signals requiring minimal storage space. Since each word is stored in its own individual memory, simple logic decoding can be used to accomplish sequencing (for a desired message) without a need for complicated programming. Even though the synthesized voice is reproduced electronically, all the voice inflections and natural qualities are preserved. The resulting synthesized voice response is so natural sounding that it is difficult to distinguish it from the original speaker.

Bell Laboratories has designed and built speech response equipment based on digital speech encoding techniques. The Adaptive Differential Pulse Code Modulation (ADPCM) method of digitizing speech offers several advantages in digital voice response applications, one being good quality speech response. Another is that the entire encoding/decoding process can be performed by inexpensive hardware in real time without the requirement of CPU time for processing. By fully exploiting the unique features of ADPCM coding, it is possible to automatically create and edit a vocabulary set. One currently used application of the voice response equipment is that of computer aided wiring by voice. Instead of using a printed wire list, the wire man works from a spoken list (a

cassette tape recording) generated automatically by a voice response system based on ADPCM coding. The main disadvantage of the digital encoding methods is the high amount of storage required. A typical vocabulary set of 100 words would require approximately 100,000 words of storage.

A Naval Research Laboratory Program to develop a practical speech synthesis system for command and control applications has resulted in a computer program that translates English texts into synthetic speech. In the present system, speech is synthesized by the VS-6 Synthesizer unit.

The Perception Technology Corporation Voice Response unit is based on pre-stored human speech. Magnetic recordings of high quality speech are sampled at 10Khz rate with six-bit resolution which results in 60,000 bits/second digitized speech. The digitized data is then compressed by a proprietary PTC developed method and stored in memory. While the technique results in high quality speech response, storage requirements (approximately 256 computer words (average) for each word of speech) are prohibitive.

Other areas of application being evaluated are voice readout for meters, calculators, hospital patient monitoring, fault warning devices, and various interactive training devices.

While the above is not an exhaustive review of the industry and government activity in voice output/input, it is quite representative of the major activity having direct relation to the goals and objectives of this study.

SECTION V

CONCLUSIONS

5.1 CONCLUSIONS

Voice data entry peripherals and voice output response peripherals have already advanced to a state of practicality. Isolated word (discrete) recognition peripherals have achieved accuracy rates of approximately 98% for selected inputs. An interactive feature precludes use of an erroneous recognition decision by prompting immediate error correction. While on-going research on the problem of connected (continuous) word recognition will someday result in an "ideal" system that fully "understands" continuous speech input, available technology makes it possible to effectively alleviate complex man-machine interface problems right now!

Most military voice communication employs a set disconnected speech format (isolated words and phrases) for command and control functions which appear to be continuous in nature (e.g., FIRE MISSION DEMO and DIAGNOSTIC TESTS). Present voice O/I capability is more than adequate for application to most military requirements.

Voice input, with proper structuring, has potential to be used as a high level input to a compiler, thereby alleviating the programming task.

Various acoustical signals within the audio band could be processed employing basic techniques used in speech signal processing. Unique applications of such systems would include acoustical fault diagnostics (e.g., M151 Engine Diagnostic).

Man-machine dialogue can be compared to instructing a subordinate in the performance of his duties. Depending on the application, the man or the machine can be the subordinate. In either case, the job gets done . . . without hesitation or misinterpretation! Voice operating systems are being successfully applied to many fields.

The final factor that will determine if speech recognition/response systems achieve wide spread application is an economic one. When voice systems cost less than the training to accomplish the equivalent functions performed by humans, a strong cost justification favoring such systems results. The potential significant reduction in training costs associated with the use of VOICE as an output/input subsystem for automatic test equipment should be sufficient stimulus to expand this activity. The activity covered by this report clearly demonstrates that VOICE has made the transition from exploratory development to advanced development. Future developments in new applications and increased capability voice systems are expected to follow.

5.2 POTENTIAL APPLICATIONS

Voice output/input is applicable to a myriad of fields. The following list, while not all-inclusive, is representative of potential application areas:

- a. Direct Material Handling and Sorting.
- b. Voice Control of Manufacturing Operations.
- c. Law Enforcement and Military Computer Communication Links.
- d. Voice Controlled Machinery, Typewriters, Calculators, etc.
- e. Quality Control Inspection Processes.
- f. Control of Physical Access Through Positive Voice Identification of Personnel.
- g. Computer Programming.
- h. Computer-based Education Systems.
- i. Voice Processing Technology Expanded to Acoustic Diagnostic Systems.
- j. A Comprehensive "Computer-based Consultant" System to be Used as a Design, Production, Inspection, and Maintenance Aid.
- k. Voice Output/Input Control of Test, Measurement, and Diagnostic Equipment.

D I S T R I B U T I O N

ARMY

Commander
U.S. Army Materiel Development
and Readiness Command
5001 Eisenhower Avenue
Alexandria, VA 22333

1 ATTN: DRCDL - J. A. Bender
1 ATTN: DRCDE-T
1 ATTN: DRCDE-DS

Commander
U.S. Army Tank Automotive
Materiel Readiness Command
Warren, MI 48090

1 ATTN: DRSTA-RGD
1 ATTN: DRSTA-MST - Mr. R. Watts

Commander
U.S. Army Aviation Systems Command
P.O. Box 209
ATTN: DRSAV-FPP
St. Louis, MO 63120

Commander
U.S. Army Electronics Command
Ft. Monmouth, NJ 07703

1 ATTN: DRSEL-SA - Mr. R. Caccamise
1 ATTN: DRSEL-TL-M- Mr. M. Tenzer
1 ATTN: DRSEL-MA-D- Mr. J. Carter

Commander
Tobyhanna Army Depot
Tobyhanna, PA 18466

1 ATTN: DRXTO-MI-O - Mr. J. Buckland
1 ATTN: Dep. Dir, Mr. W. Morris

Commander
U.S. Army Materiel Systems
Analysis Agency
ATTN: DRXSY-CC - Mr. D. Lynch
Aberdeen Proving Ground, MD 21005

Commander
U.S. Army Missile Command
Huntsville, AL 35809

1 ATTN: DRSMI-MA
1 ATTN: DRSMI-RLD - Mr. D. Thurman
1 ATTN: DRSMI-RLE - Mr. G. Hubbard
1 ATTN: Director of Maintenance, TMDE Ofc

Commander
Army Satellite Communications Command
ATTN: DRCPM-SC-8B
Ft. Monmouth, NJ 07703

Commander
U.S. Army Maintenance Management Center
ATTN: DRXMD-TR - Mr. W. Stratton
Lexington, KY 40507

Commander
U.S. Army ARADMAC
ATTN: SAVAE-EMP - Mr. L. Baca
Corpus Christi, TX 78419

Army Security Agency
ATTN: IALOG-RME
Arlington, VA 22212

Commander
Sacramento Army Depot
Sacramento, CA 95813

1 ATTN: DR SXA-MME - Mr. H. Kuyama
1 ATTN: Dep. Dir, Maint - Mr. A. Weaver

Commander
Picatinny Arsenal
ATTN: TSD - Mr. D. Morlock (Bldg 352)
Dover, NJ 07801

Commander
U.S. Army Logistics Center
ATTN: ATCL-MC - Mr. C. Adenauer
Ft. Lee, VA 23801

D I S T R I B U T I O N

ARMY (continued)

Commander
TRASANA
ATTN: ATAA-TDL
White Sands Missile Range, NM 88002

DA-ODCSLOG
ATTN: DALO-SMM-E - Mr. Nichols
Washington, DC 20310

COL E. A. Viereck, Jr.
ASA L&L
Room 3E619
Pentagon Building
Washington, DC 20310

PM, ATSS
ATTN: DRCPM-ATSS, LTC J. Gabrysiak
Ft. Monmouth, NJ 07703

Commander
U.S. Army Troop Support Command
4300 Goodfellow Boulevard
St. Louis, MO 63120

Commander
U.S. Army Armament Command
ATTN: DRSAR-RDG
Rock Island, IL 61201

Headquarters, DA
DAMA-CSS
Washington, DC 20310

PM, ARTADS
ATTN: DRCPM-TDS-TF - Mr. David Blank
Ft. Monmouth, NJ 07703

Commander
U.S. Army TARADCOM
ATTN: DRDTA-RGD - Mr. J. Steyaert
28251 Van Dyke
Warren, MI 48090

D I S T R I B U T I O N

NAVY

Headquarters
Naval Materiel Command
ATTN: Director, Automatic Test
Equipment Management & Technology
Office (MAT 036T)
Washington, DC 20360

Commander
Naval Air Systems Command
ATTN: AIR 53424, Mr. M. Myles
Washington, DC 20360

Commander
Naval Electronic Systems
Engineering Center
ATTN: 03T, Mr. F. Fernandez
P.O. Box 80337
San Diego, CA 92138

Commander
Naval Electronic Systems Command
ATTN: ELEX 4802, Mr. G. Margulies
Washington, DC 20360

Commander
Naval Sea Systems Command
ATTN: SEA 9822B, Mr. John Yaroma
Washington, DC 20360

Commander
Naval Electronics Laboratory Center
ATTN: Code 4050, Mr. D. Douglas
271 Cataline Boulevard
San Diego, CA 92152

D I S T R I B U T I O N

AIR FORCE

Commander
Kelly Air Force Base
San Antonio, TX 78241

1 ATTN: SA-ALC/MMIN-Mr. J. Ferrell
1 ATTN: SA-ALC/XRXM-Mr. T. Daily
1 ATTN: SA-ALC/MMEC-Mr. J. Cotnam
1 ATTN: SA/ALC/MAGT-Mr. H. Arant

Headquarters
AFSC/LGM
Andrews Air Force Base
ATTN: Mr. Clark Walker
Washington, DC 20334

Headquarters
USAF/LGYE
ATTN: Mr. C. Houk
Washington, DC 20330

Commander
Wright Patterson Air Force Base
Dayton, OH 45433

1 ATTN: ASD/FNEGE-Mr. D. Behymer
1 ATTN: ASD/AEG

D I S T R I B U T I O N

OTHER

Defense Documentation Center (12)
Cameron Station
Alexandria, VA 22314

D I S T R I B U T I O N

FRANKFORD ARSENAL

Commander
Frankford Arsenal
Philadelphia, PA 19137

1 ATIN: CRF 107/1
1 ATIN: TD 107/1
1 ATIN: PA 107/2
1 ATIN: GC, 28/1
1 ATIN: QAA-R 119/2
1 ATIN: PP, Mr. G. White 64/4
2 ATIN: TSP-L 51/2
1 ATIN: FCF-C 202/1
40 ATIN: FCF 201/1
1 ATIN: FCF-S 201/1
1 ATIN: FCF-F 201/1
6 ATIN: TSP-T 51/2