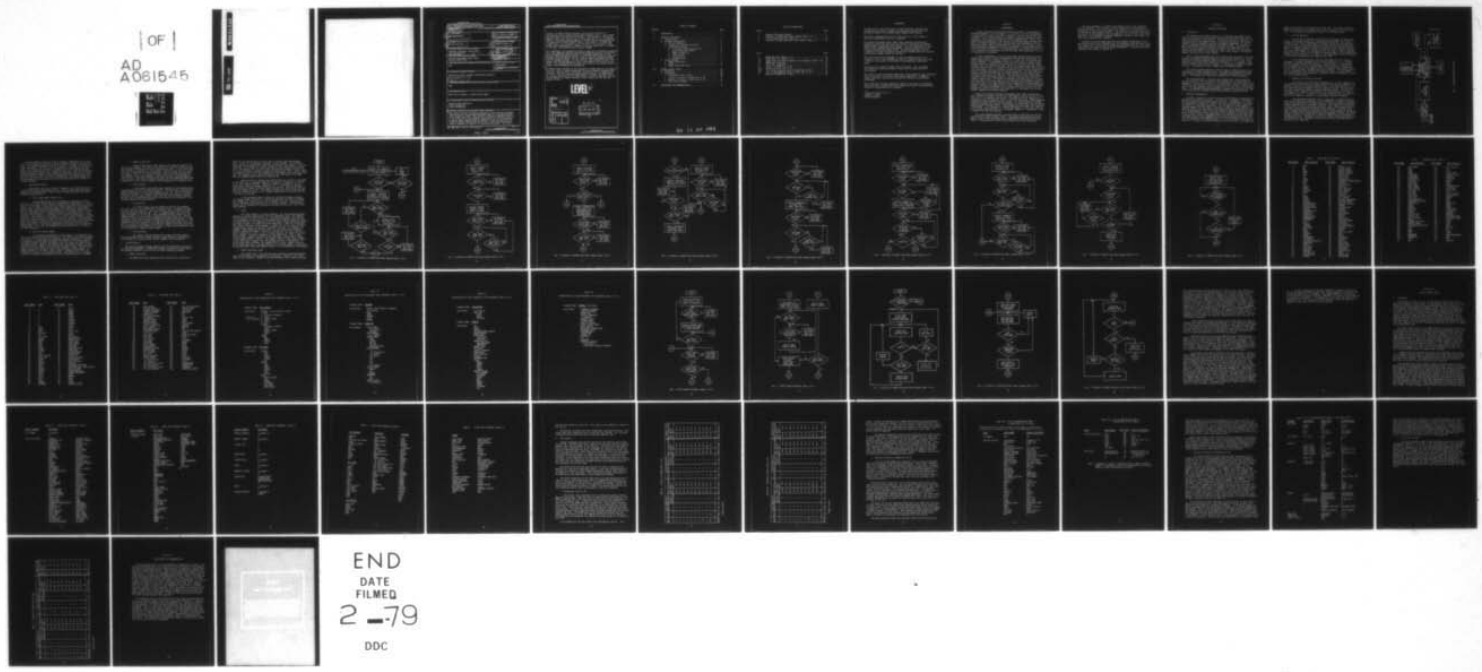
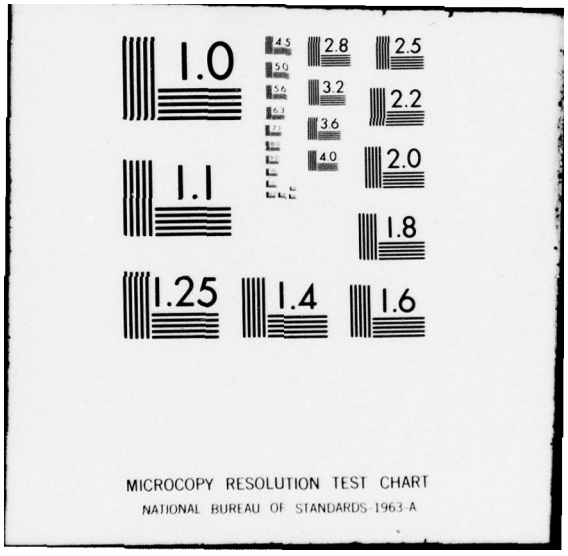
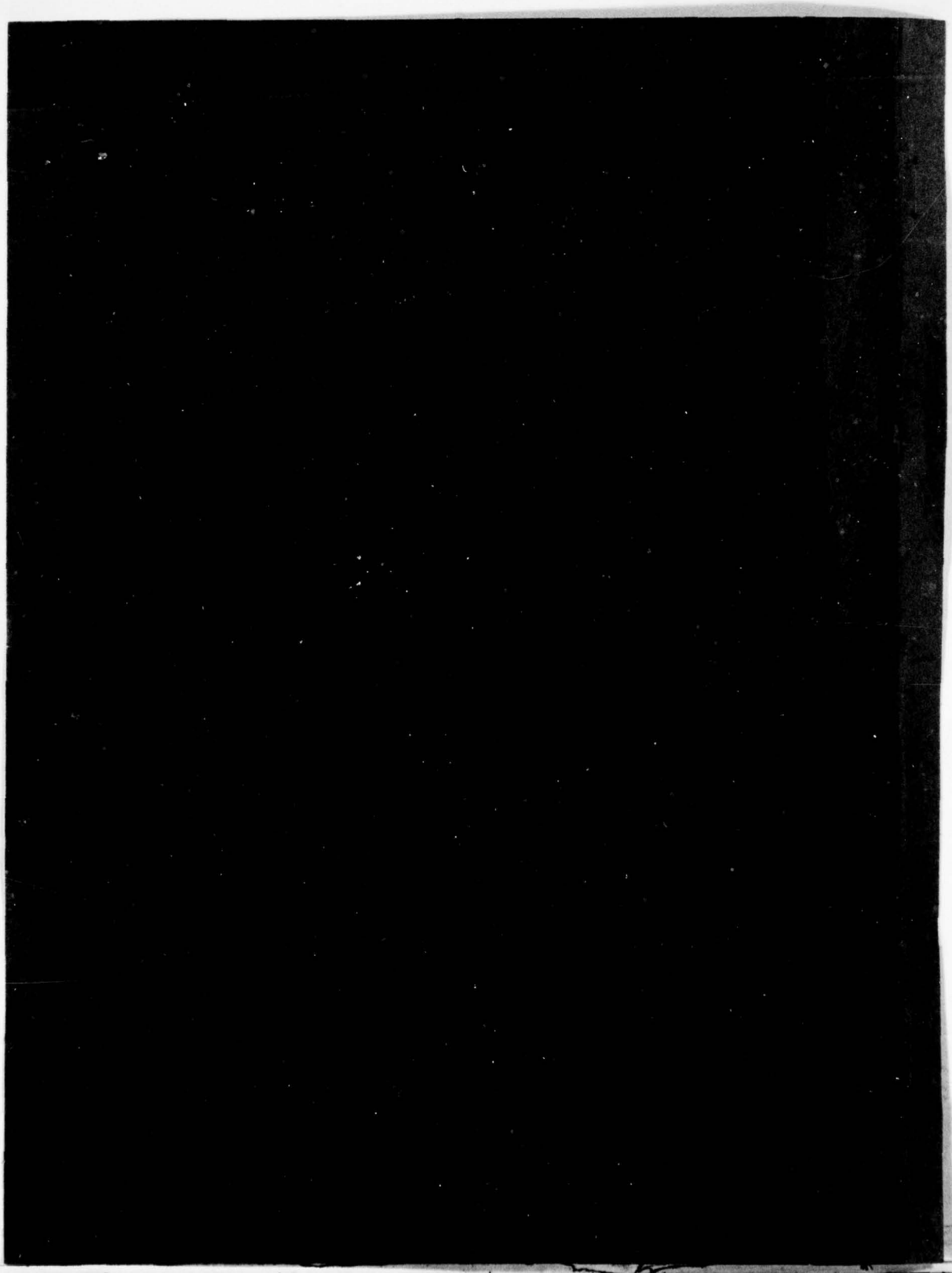
AD-A061 545 THRESHOLD TECHNOLOGY INC DELRAN N J F/G 5/8
WORD RECOGNITION.(U)
SEP 78 P B SCOTT
UNCLASSIFIED RADC-TR-78-209 F30602-76-C-0394 NL



END
DATE FILMED
2-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: RADC-TR-78-209	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): WORD RECOGNITION	5. TYPE OF REPORT & PERIOD COVERED: Final Technical Report
	6. PERFORMING ORG. REPORT NUMBER: N/A
7. AUTHOR(s): Phillips B. Scott	8. CONTRACT OR GRANT NUMBER(s): F30602-76-C-0394
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Threshold Technology, Incorporated, 1829 Underwood Blvd, Delran NJ 08075	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 63701B, 32010316, 03
11. CONTROLLING OFFICE NAME AND ADDRESS: Rome Air Development Center (IRAA), Griffiss AFB NY 13441	12. REPORT DATE: September 1978
	13. NUMBER OF PAGES: 53
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Same	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

12 57 p.

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

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: Richard Vonusa (IRAA)

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

Automatic Word Recognition
Speech Recognition
Pattern Recognition

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

This report describes the design, principles of operation and performance characteristics of an Advanced Development Model of a limited-vocabulary voice recognition system. This system is intended to demonstrate the use of voice data entry in cartographic applications such as the processing of Digital Radar Landmass Simulator (DRLMS) data and the preparation of Flight Information Publications (FLIPS) data. Three separate operating computer programs were supplied for system demonstrations. The first

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

390 798

program, the DRLMS Fixed Node program, was constructed expressly on instructions from the Defense Mapping Agency Aerospace Center. This program has fixed nodes with conditional branching. It is designed to simulate the manner in which voice could be used to input DRLMS FAC data to a cartographic computer. The second program simulates a possible mode of inputting data for the construction of FLIPS terminal charts. The program is a variable node program with a total vocabulary size of 500 words. This structure of this program allows recognition of a maximum of 100 words in any node of a maximum of 30 nodes. Node structures may be easily changed.

This cartographic word recognition system is based upon the Threshold Technology Inc. (TTI) commercial VIP-100 isolated-word recognition system. The VIP-100 can be automatically adapted on-line for individual speakers and/or words. The principal speech recognition modules of the VIP-100 are a speech preprocessor designed and manufactured by TTI and a general purpose minicomputer running with TTI designed software. For this contract, RADC furnished as GFE, a Data General Nova 800 minicomputer with 16K of core memory to be included in the system. TTI furnished additional memory plus ancillary input/output devices to enhance system performance and capabilities.

Performance tests conducted at TTI disclosed an average system word recognition accuracy of 99% for talkers experienced with word recognition systems. These tests were conducted with two different vocabularies, one for DRLMS and one for FLIPS. Each test series was performed by 10 talkers. The DRLMS talker set included eight male and two female talkers, the FLIPS talker set nine male and one female. The lowest individual talker accuracy was over 95%, the highest was 100%.

LEVEL II

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

EVALUATION

The objective of this effort was to develop and apply existing voice recognition technology to enhance the cartographic digitization and editing processes at the Defense Mapping Agency (DMA) centers.

This effort demonstrated that voice data entry is a viable means of inputting cartographic data into a computer.

An Advanced Development Model (ADM) large vocabulary isolated word recognition system was fabricated and tested by simulating two Defense Mapping Agency cartographic applications. The first application is the Digital Radar Landmass Simulator (DRLMS) in which an analyst fills out a Feature Analysis Data Table (FADT). The second application is the preparation of Flight Information Publications (FLIPS) data which involves constructing terminal charts for aviation.

Both applications were programmed to input cartographic data into a computer by voice, thus allowing the analyst's hands and eyes to be free so that he could devote his attention on performing his specialized tasks.

Voice entry was accurate (greater than 95 percent), fast, and easier than using the conventional means of data entry by paper and pencil and keypunch.

As a result of this successful simulation, future plans call for a follow-on effort to procure a voice data entry system, interface it to DMAAC's digital data base, and test and evaluate the system in an operational environment.

This technology will have significant impact on the speed to efficiently enter data into a computer and will eliminate the need to acquire skilled typists for interaction with a computer.

Richard S. Vonusa

RICHARD S. VONUSA
Project Engineer

Section I

INTRODUCTION

A major limitation of most data processing systems is the problem of getting information into the computer. As a result, more and more emphasis is being placed on source data entry techniques. For on-line systems, capturing data at the source requires continuous interaction between a man and a machine. This interaction implies a two-way communication. First, the human provides information to the system, and then the system provides feedback to the human indicating that the last input was understood as well as supplying results and/or an indication that an action will be taken. Each time a man-machine communication system is designed, a careful analysis must be made of what information is to be transmitted in each direction, how it is to be transmitted or displayed, and when and under what conditions it is to be transmitted or displayed.

An area of importance to the Air Force where enhanced data entry techniques can be applied involves cartographic digitization and editing at the Defense Mapping Agency Aerospace Center in St. Louis, MO. with present data entry techniques. A cartographic analyst now must prepare a Digital Radar Landmass Simulator (DRLMS) culture manuscript by the use of overlays, handwritten notes, and other manual aids. He or she transfers data for each FAC (feature analysis code) to a Feature Analysis Data Table (FADT) which consists of as many pages as may be required to record all pertinent data necessary to describe the areas portrayed on the feature manuscript. The FADT table has numerous columns into which must be entered by pen or pencil number codes describing the features for each FAC. The FADT data is finally converted to a digital record by keypunch operators who must transcribe the FADT onto a computer digital tape. Any method to automate this processing will increase the efficiency of the data input and will decrease the likelihood of errors. Voice data input is the most natural form of communication between humans and machines because vocal communications is most frequently used between humans. However, the obvious advantages of voice data entry can be lost if the word recognition system used places such constraints on the humans that the man-machine communication becomes unnatural.

During this program an Advance Development Model (ADM) isolated-word Voice Recognition System (VRS) has been developed to explore the use of voice data entry in cartographic processing. This ADM-VRS has the capability of recognizing up to 500 separate words in syntactic structures. Three structured programs have been supplied with the system. The first program, the DRLMS Fixed Node program, was constructed expressly on instructions from the Defense Mapping Agency Aerospace Center. This program has fixed nodes with conditional branching. It is designed to simulate the manner in which voice could be used to input DRLMS FAC data to a cartographic computer. Interfacing to such a computer has not been provided.

The second program simulates a possible mode of inputting data for the construction of FLIPS terminal charts. The program construction was based upon available FLIPS information and by agreement between RADC and TTI personnel. The program has a fixed node structure and a fixed vocabulary.

The third program is a variable node program with a total vocabulary size of 500 words. This structure of this program allows recognition of a maximum of 100 words in any one node of the structure. Up to 30 nodes can be included in the sentence structure. The words to be included in each node as well as the node sequences can be changed at will by the use of the Teletype input. All programs are speaker dependent; therefore, they require the input of training data by a speaker.

Section II of this report describes the hardware configuration of the system and outlines with flowcharts the three programs which make up the system software set. Section III describes the preliminary and the final system tests which were used to evaluate system performance. Conclusion and recommendations are covered in Section IV.

Section II

TECHNICAL DISCUSSION

A. Introduction

The hardware and software features of the Advanced Development Model (ADM) isolated-word Voice Recognition System (VRS) developed under this contract are described in this section. This VRS system has the capability of recognizing up to 500 separate words in syntactic structures. The system hardware includes equipment supplied by Threshold Technology Inc. (TTI) as well as government furnished property furnished by RADC. The hardware is based upon the VIP-100 word recognition system sold commercially by TTI. Several ancillary hardware items not included in the basic VIP-100 configuration have been added to this ADM-VRS. These items include a GFP digital magnetic tape unit and a video terminal.

Three structured programs have been supplied with the system. The first program, the DRLMS Fixed Node program, was constructed expressly on instructions from the Defense Mapping Agency Aerospace Center in St. Louis. This program has fixed nodes with conditional branching. It is designed to operate in a manner which would allow input to a cartographic computer of DRLMS FAC data. Interfacing to such a computer has not been provided.

The second program simulates a possible mode of inputting data for the construction of FLIPS terminal charts. The program construction was based upon available FLIPS information and by agreement between RADC and TTI personnel. The program has a fixed node structure and a fixed vocabulary.

The third program is a variable node program with a total vocabulary size of 500 words. This structure of this program allows recognition of a maximum of 100 words in any one node of the structure. Up to 30 nodes can be included in the sentence structure. The words to be included in each node as well as the node sequences can be changed at will by the use of the Teletype input. All programs are speaker dependent; therefore, they require the input of training data by a speaker before the inputting of operational data.

B. Hardware Configuration

The VRS is basically a VIP-100 isolated word recognition system which is described very briefly as follows. The VIP-100 is an adaptive system which must be trained for individual talkers and words. Consequently, the system is automatically adjusted or "tuned" to the voice characteristics of different users in a very short time period. By the input of a small number of training samples into the device to provide a reference set of features, the decision criteria for each word in the vocabulary can be modified or trained in an optimum manner. Thus, the system stores in memory an individual reference set of word features for each word in the vocabulary and for each talker in the system. After a cartographer trains the system new words spoken during normal operation are compared with the stored references and a "closest fit" is selected as the recognized word. It is also possible to obtain a "no decision", or reject, when the characteristics of all words in the reference

memory are dissimilar to the particular input word. This reject capability effectively prevents false recognition of words outside the selected vocabulary. It also prevents recognition response to extraneous noises.

1. VIP-100 Operation

The VIP-100 has two modes of operation with regard to speech recognition, training and recognition. During the training mode, the VIP-100 automatically extracts a time-normalized feature matrix for each repetition of a given word. A consistent matrix of feature occurrences (between repetitions) is required before the features are stored in the reference pattern memory. A template threshold factor is chosen such that a feature occurrence (in a given time segment) is considered valid only when it occurs a minimum number of times relative to the number of training samples. The hardware and software have the flexibility to allow any one vocabulary word to be trained or for the entire vocabulary to be trained at one time.

In the recognition mode, each new word spoken into the system is processed in a manner analagous to the training procedure - i.e., features extracted, digitized and time normalized. The resultant test word matrix then is compared digitally to each stored reference matrix. Similarities and dissimilarities in each compared matrix are appropriately weighted and the net result provides a weighted correlation product. Correlation products also are generated after shifting the input word matrix ± 1 time segment. The stored reference word producing the highest overall correlation is selected as the recognized word.

2. Hardware and Software Processors

The essential word recognition elements of a VIP-100 in general and this VRS-ADM in particular are a speech preprocessor and a minicomputer. In the VRS, as shown in the block diagram of Fig. 1, the preprocessor is a TTI model 8040 and the minicomputer is a Data General Nova 800. The 8040 serves to extract significant acoustic features from input speech. The Nova 800 minicomputer then compares these features with stored features for words in the vocabulary in use. A decision as to which vocabulary word was spoken is made by the computer according to a predetermined decision algorithm.

The preprocessor operates using a combination of hardwired digital logic and analog signal processing. The minicomputer is software controlled. The features extracted by the preprocessor are a selected subset (including complex combinations) of 32 acoustic features. Each feature is extracted by a combination of analog operations and binary logic. These features are of two types, primary features and phonetic-event features. Features of the former category describe the spectrum directly by indicating local maxima and areas of increasing or decreasing energy with frequency (slopes). The latter category consists of features which represent measurements corresponding to phoneme-like events. Included in this set are vowels, nasals and fricatives. The preprocessor also determines accurately in hardware the beginning and ending points of each word.

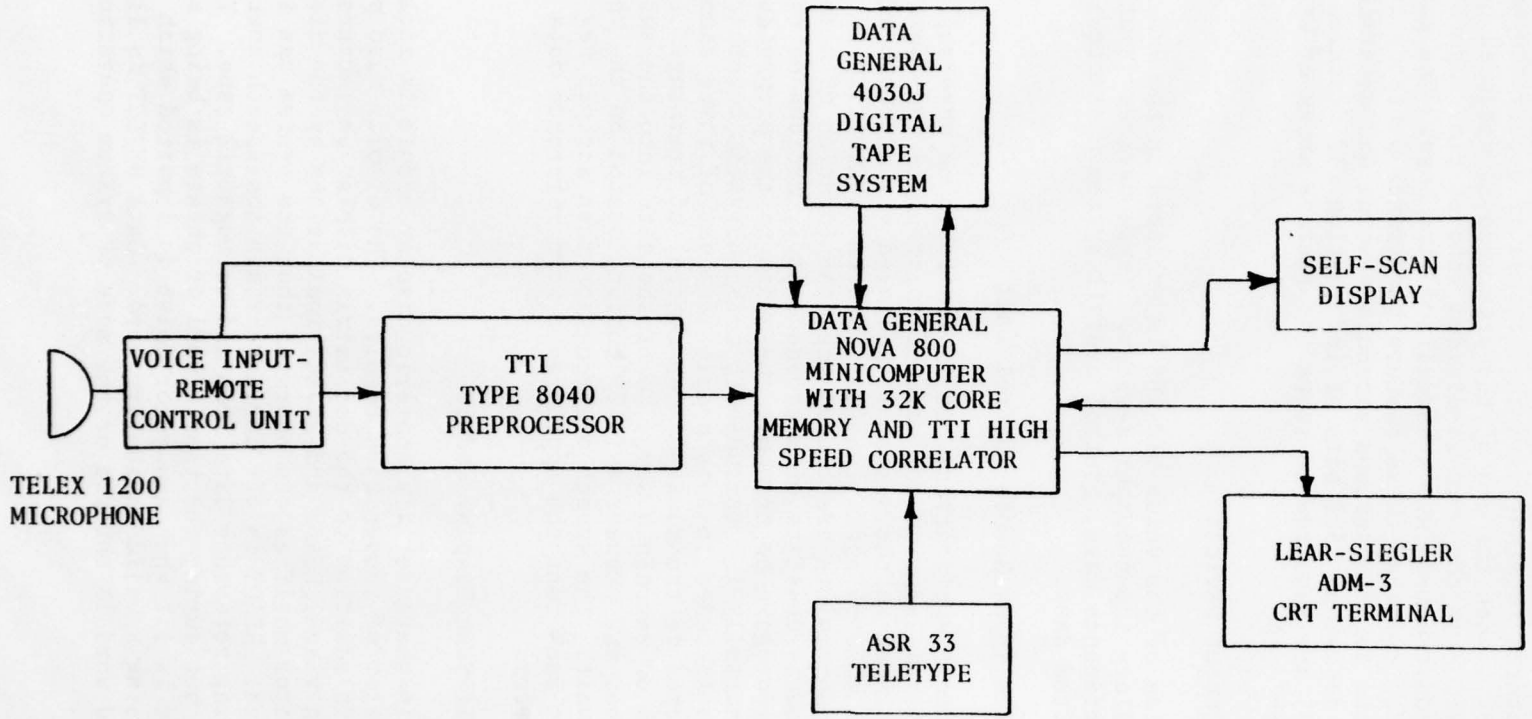


Fig. 1 Complete VRS Block Diagram.

For each spoken word the 32 acoustic features, represented in a binary form, are continuously inputted to the minicomputer through a special interface. These features together with their times of occurrence are stored in a short-term memory. When the end of the utterance is detected by the preprocessor, the duration of the word is divided into 16 time segments and the features are reconstructed into a normalized time base. The pattern-matching logic subsequently compares these feature occurrence patterns to the stored reference patterns for the various vocabulary words and determines the "best fit" for a word decision. 512 bits of information (32 features mapped into 16 time segments) are required to store the feature array of an utterance or reference pattern.

3. Input-Output Devices

Operation of the voice VIP-100 is enhanced in the VRS-ADM by the use of several ancillary input-output devices. These devices input the speech signal, store reference data, control overall system operations and display previously inputted data.

a. Voice Input-Remote Control Unit

In most applications, it is desirable to physically locate a VIP-100 in a central location for ease of maintenance and logistics. Therefore, the voice input to the system is achieved via a remote audio and control module. In this mode of operation, more system flexibility is obtained because minimum constraints are placed upon the location of a small remote module. This module consists of microphone jack, microphone equalizer circuitry, and sufficient gain to transmit the audio to the preprocessor. A gain control switch normally is included in this subsystem together with a level indicator. Also included in this module is a set of three thumbwheel switches which are connected to the Nova 800 for control of training. If an operator desires to train or retrain a word, he or she sets into the switches the appropriate word number and presses a TRAIN button located on the Voice Input/ Remote Control unit. The speech preprocessor then accepts new training data and processes it such that the appropriate word reference data is stored in the Nova 800 memory.

b. Self-Scan Display Module

A 16-character alphanumeric display module is included for display to an operator of recognition results. The display also provides prompts for training. In addition to the dot matrix display characters, the operator is provided with visual READY and REJECT indications by the display module. The READY indicator notifies the operator that the system can input and process speaker data. READY is off during certain non-speech operations such as reading or writing reference data from and to magnetic tape. The READY indicator also goes out during the time a word or phrase is being spoken. The REJECT indicator is lit when a word or noise is inputted which cannot be recognized as being a valid vocabulary word. Once REJECT is lit it remains on until a valid word is entered or the mode of system operation is changed.

c. Magnetic Tape Unit

A Data General model 4030J digital tape system was supplied as GFP for this program. This tape system provides a convenient means for long term storage of reference data for a large operator base. At any time reference data for one operator is core resident. Reference data for any vocabulary can be written to or read from magentic tape in seconds when operators are changed. Also, it is possible to store each of the three programs constructed for the VRS-ADM on magnetic tape and, therefore, change programs in a few seconds. Node plans and display messages for the DRLMS Adjustable Node program can also be stored on magentic tape independently of reference data.

d. Video Terminal

A Lear Siegler model ADM-3 video terminal has been supplied with the system to be used as a bulk data display. This unit has a 1920 character display capability, 24 rows of 80 columns, on a 12" CRT. Each of three programs uses the ADM-3 in a somewhat different manner, but the video terminal is used to display large fields of data which has been previously inputted. The ADM-3 has a standard keyboard which can be used as an alternative to the Teletype keyboard for system control.

e. Teletype Unit

A 33 ASR Teletype unit was supplied as GFP for the program. This unit serves as the primary means of system control, although the ADM-3 can serve the same function. Also, the 33 ASR is the only means for inputting new programs to the Nova 800. This is accomplished by the use of the paper tape reader of the 33 ASR. Once a program has been entered into the computer memory it can be stored on magnetic tape for future use. However, it must be entered initially by the paper tape reader of the 33 ASR. A high speed paper tape reader interface was added to the GFP Nova 800 by TTI. This interface will allow connection of a high speed paper tape reader to the system if one should be available in the future. Reference data and node plans for the adjustable node program can be stored on paper tape through the 33 ASR as an alternative to magentic tape storage.

f. System Housing

The complete VRS-ADM except for Teletype, Self-Scan display, Voice-Input/Remote Control module and ADM-3 video terminal is mounted in a 70 inch high by 19 inch wide rack cabinet complete with cooling fans.

C. System Software

The three operating programs supplied with the VRS-ADM are described in the following paragraphs. Included are complete flowcharts of each program and complete vocabulary lists for the two fixed vocabulary programs.

1. DRLMS (Fixed Node)

The DRLMS fixed node program has been constructed as specified by

DMAAC for the demonstration of voice input of digital radar cartographic data. Figure 2 is the complete flowchart of this program. This flowchart shows the display prompts and conditional branching of the program. The total vocabulary is 178 words and is fixed insofar as words displayed to the operator. The operator may, of course, make substitutions during training if he or she is consistent in usage of substitutes during data input. Table I shows the complete word list for this program with word numbers to be used for training. The operator at any time may delete the last word spoken with ERASE. The voice command CANCEL deletes the complete FAC.

After each FAC is complete the inputted data for the FAC is displayed on the ADM-3 CRT. Previously inputted FAC data may be displayed on this CRT by the voice command "REVIEW". When this command is recognized, the self-scan display shows the message "REVIEW FAC #". At this time, the operator should enter the number of the previously entered FAC. The data for that FAC then will appear on the CRT. The system will then wait for new data for that FAC to be entered by voice or for a CANCEL command which will allow data for a new FAC to be entered.

If the operator wishes to review all FAC data previously entered, the display prompt "REVIEW FAC #" should be answered by a spoken GO command. This command will cause the first FAC to appear on the CRT. Successive FACs will be displayed by pressing any key on the CRT keyboard. This reviewing process may be terminated by pressing the control key and the "A" key on the ADM-3 keyboard.

2. FLIPS

The FLIPS program is a fixed vocabulary, 150 word branching program designed to simulate inputting by voice of cartographic data. The complete FLIPS vocabulary is shown in Table II. The type of vocabulary and structure of this program could be used in constructing FLIPS terminal charts for aviation applications. The vocabulary is divided into six major categories (nodes). Each major category is divided into minor categories. Table III illustrates the category structure of the program. In many, but not all cases, the minor categories have sets of valid words associated with each. For example, the category "FILE CONTROL" as shown in Table III, has three minor categories - DRAWING, SCRATCH, AND SYMBOL. Each minor category has six, five, and four words, respectively, valid. To go to a major category, the name of that category is spoken. The minor categories are then made valid. After a minor category name is spoken, the words within that minor category may be spoken or the command "NEXT" is spoken. This command allows the next minor category (within that major category) that the operator chooses, to be entered. To go to another major category, the command "EXIT" is used. ERASE is used to delete individual words. The data inputted to the FLIPS program is displayed on the ADM-3 after each NEXT command. Figure 3 is the flowchart of this program.

3. DRLMS (Adjustable Node)

This program has a node structure and vocabulary (500 words maximum) that, within limits, may be changed by the operator. Figure 4 is the flowchart for this program. Changes in the node structure and of display messages

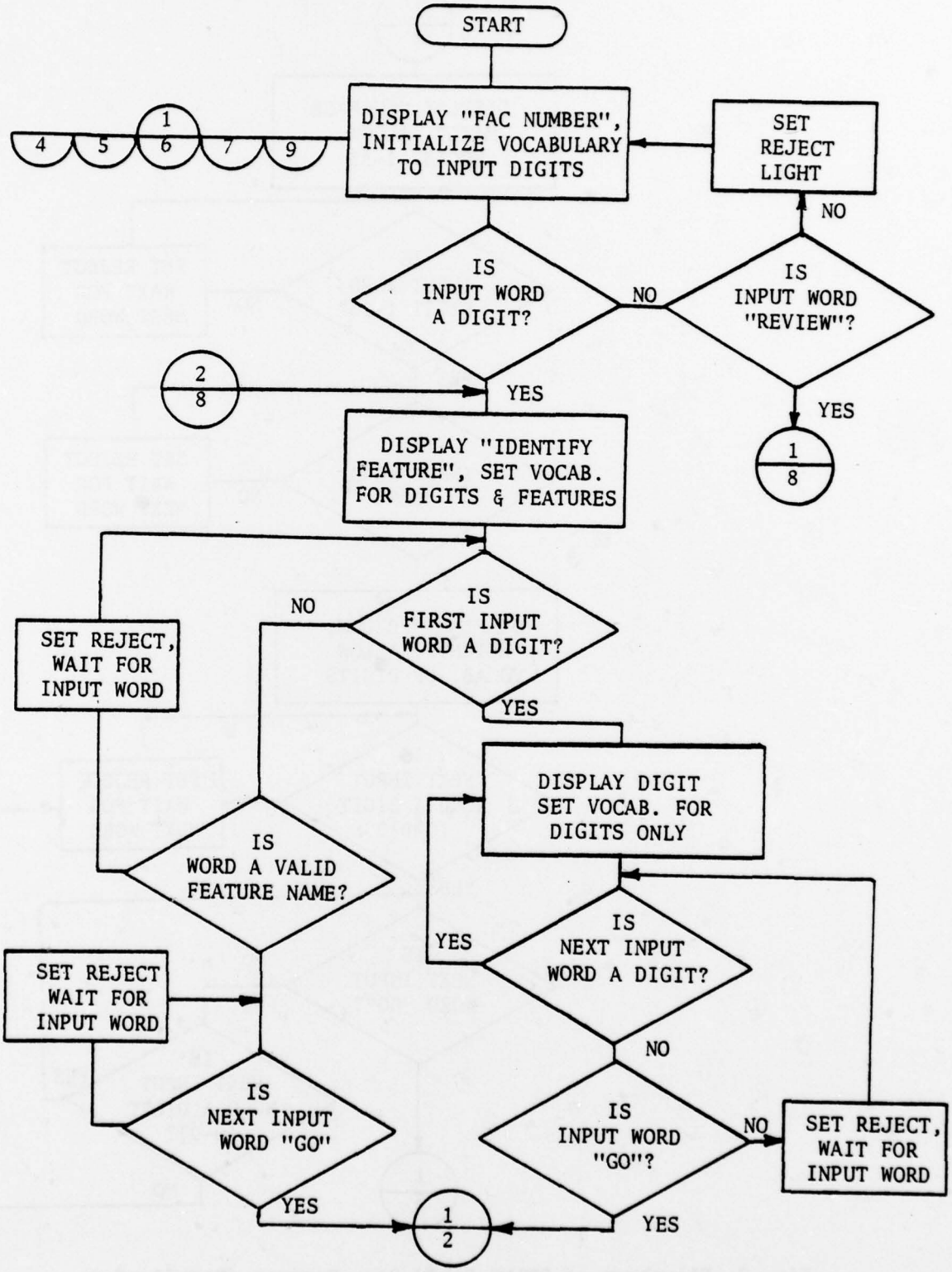


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 1 of 9)

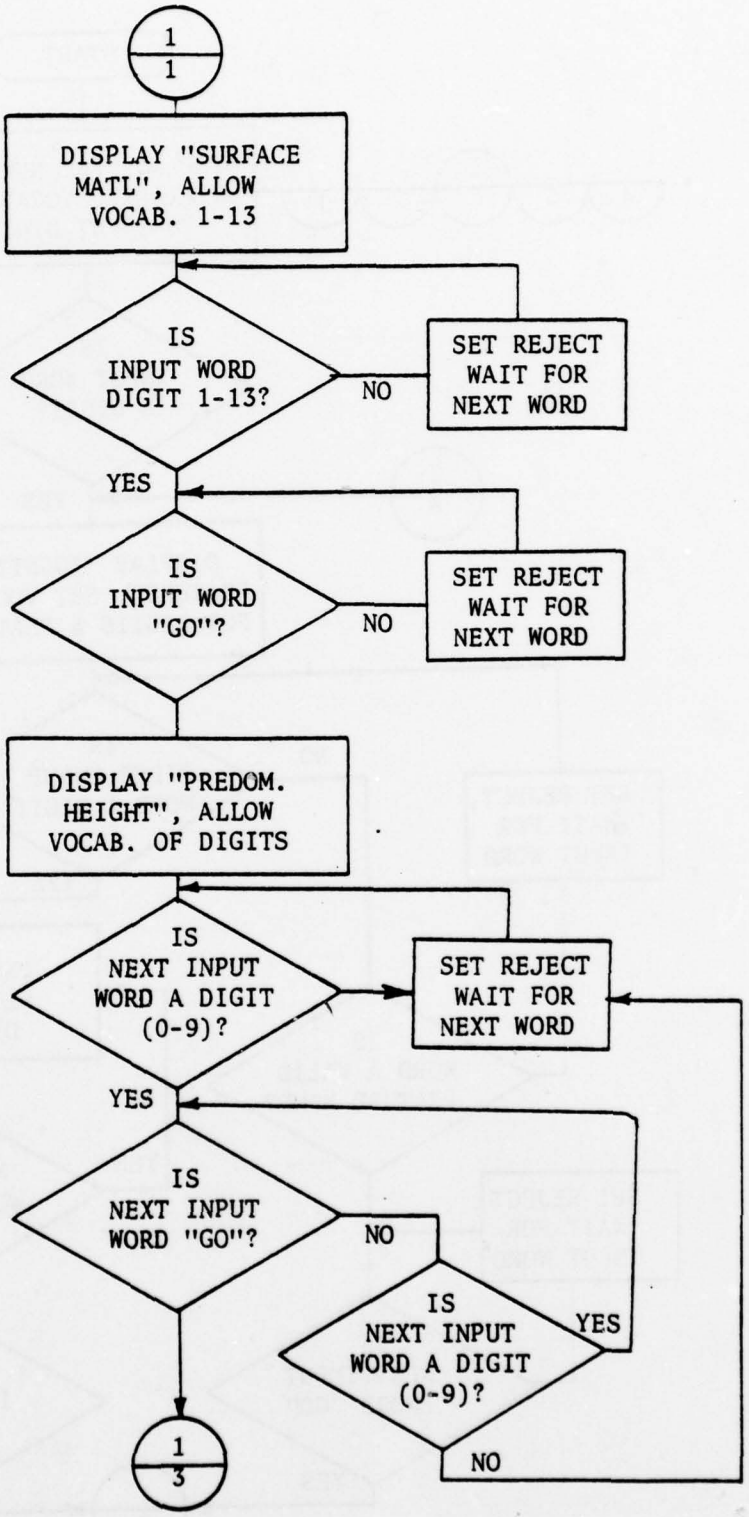


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 2 of 9)

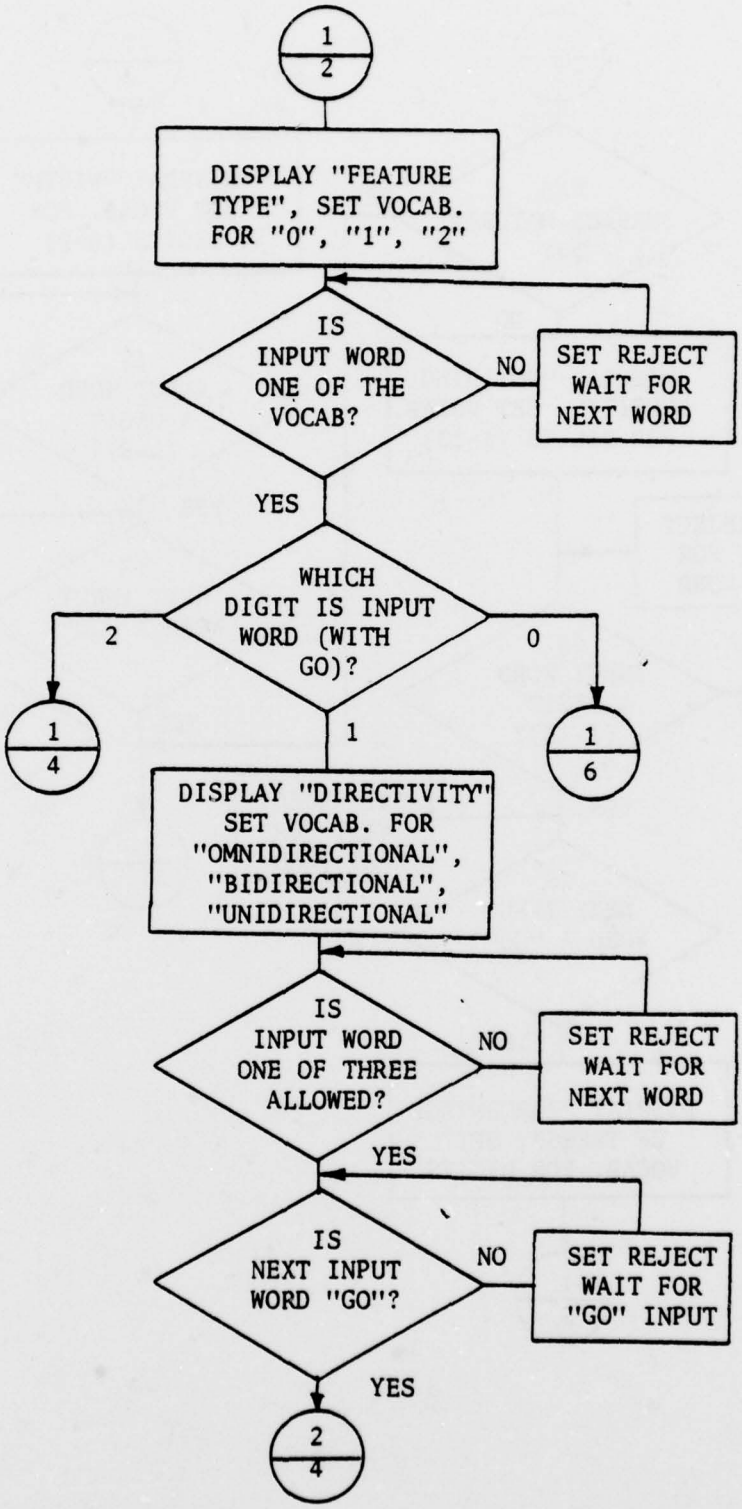


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 3 of 9)

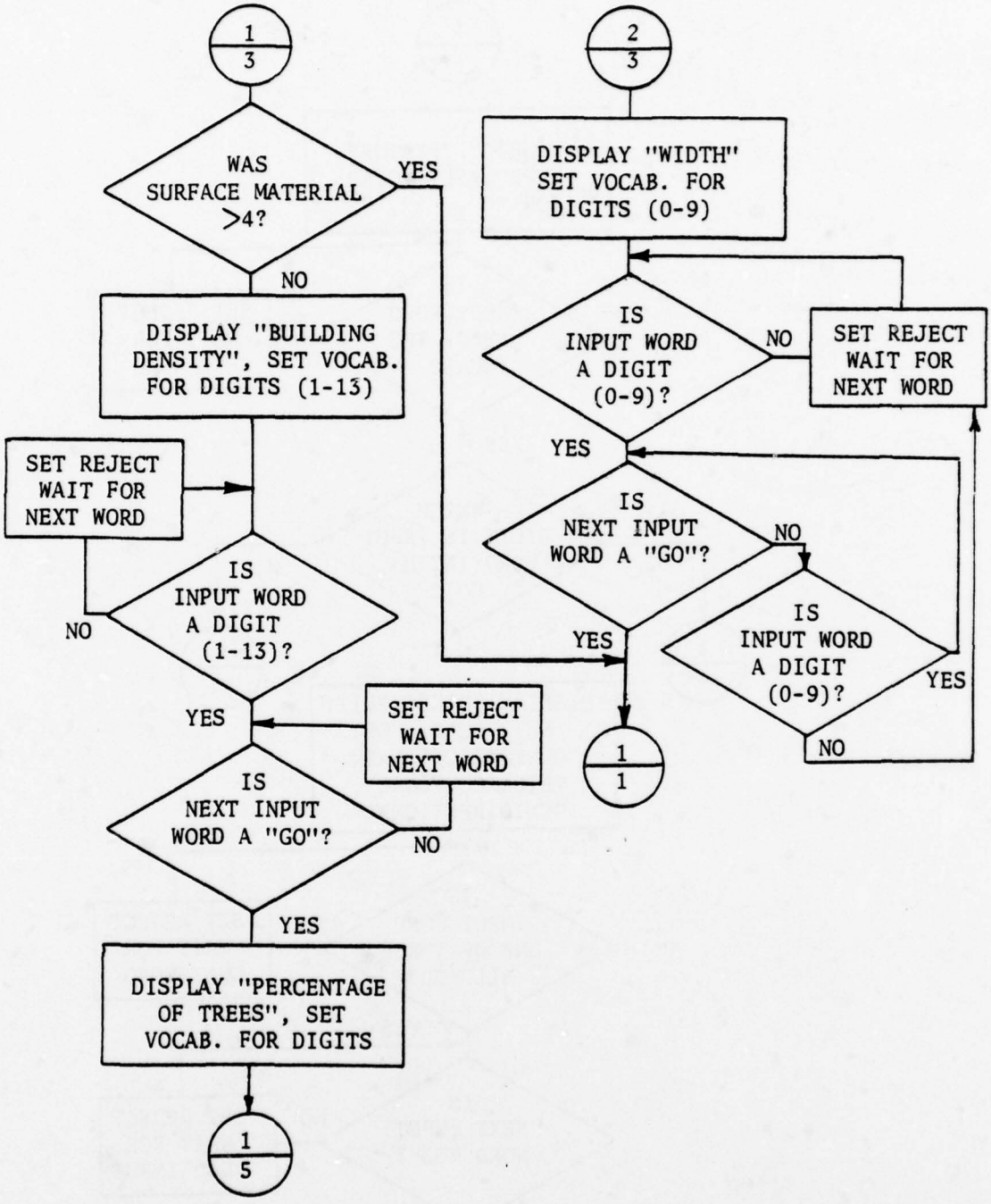


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 4 of 9)

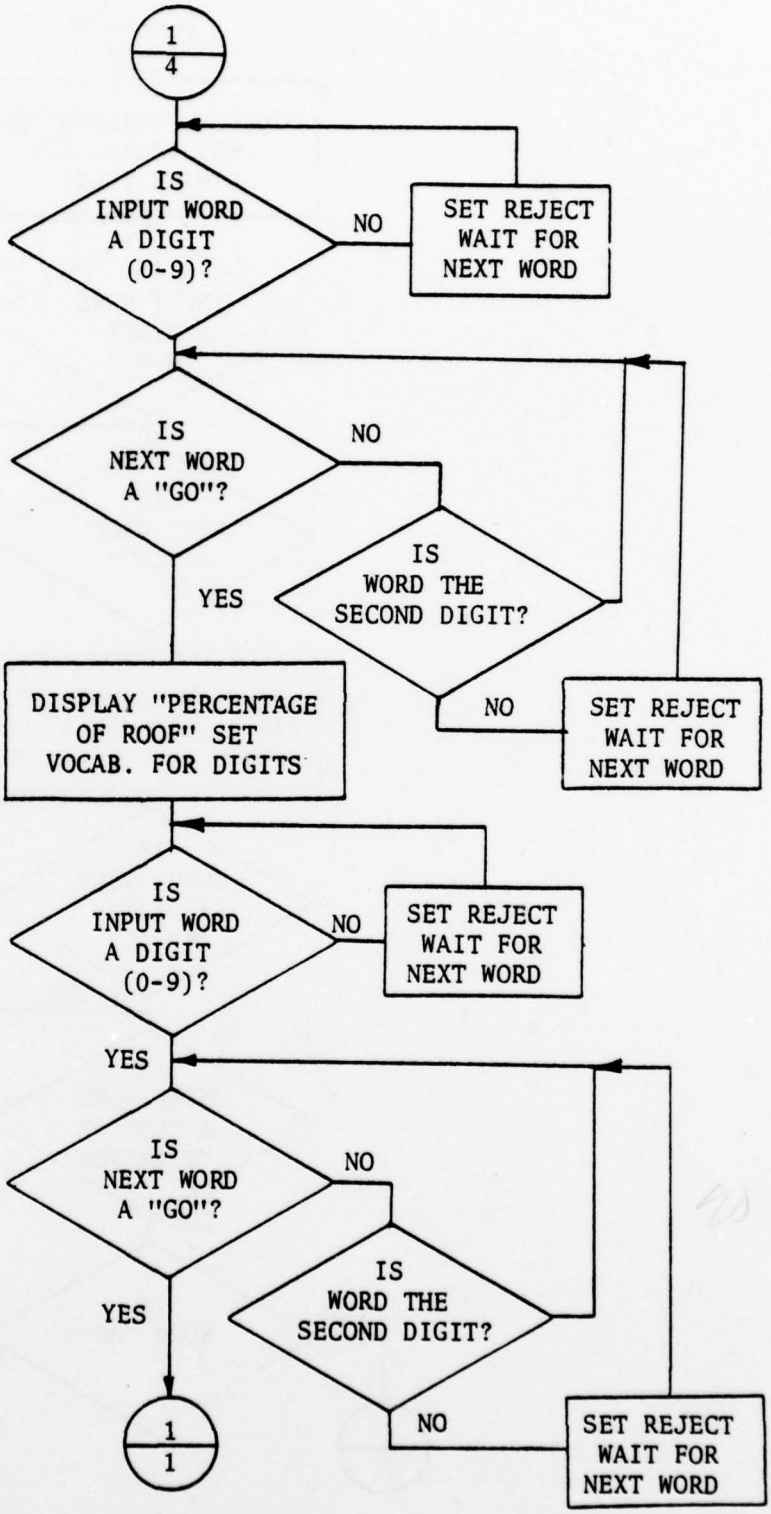


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 5 of 9)

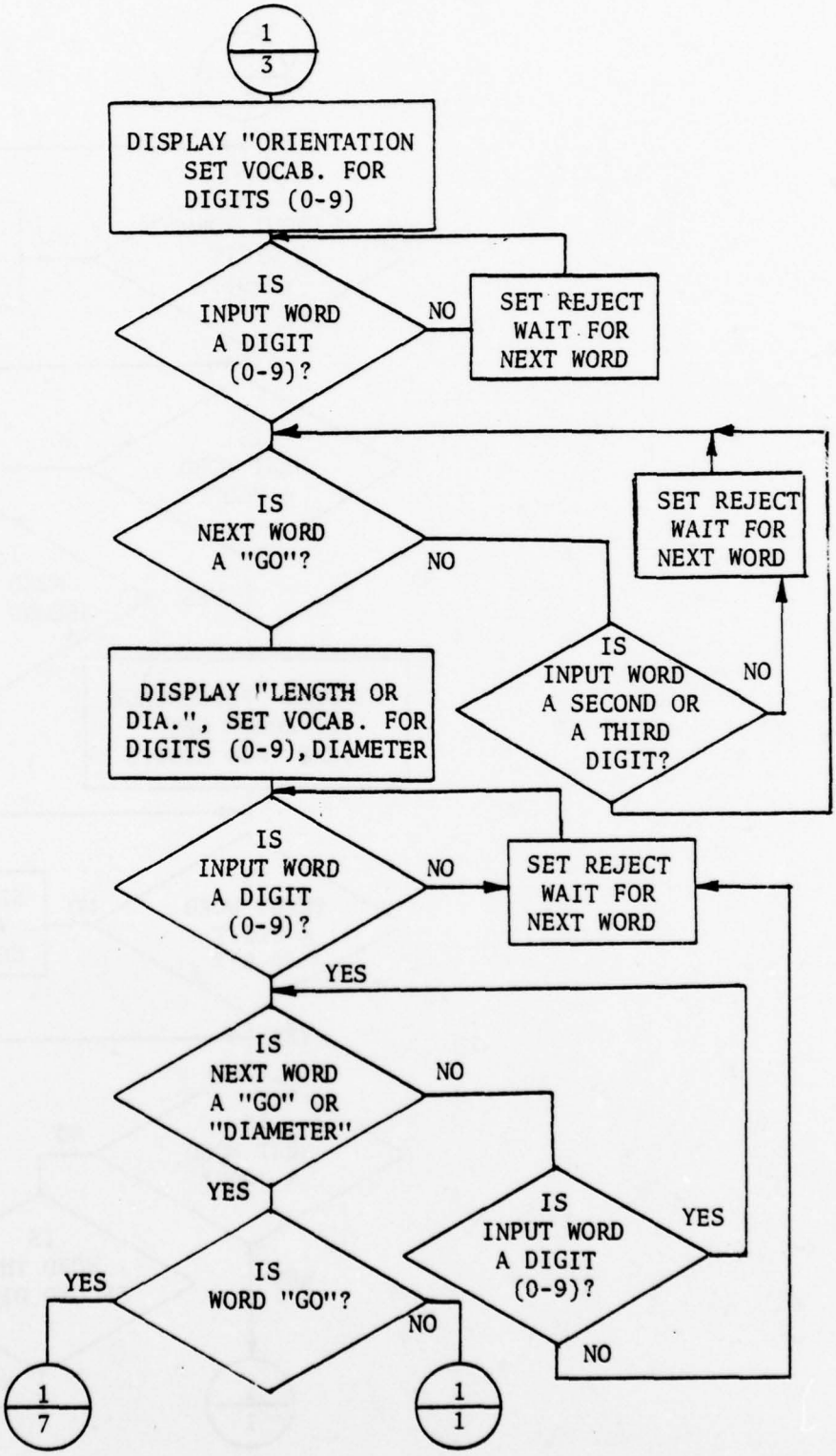


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 6 of 9)

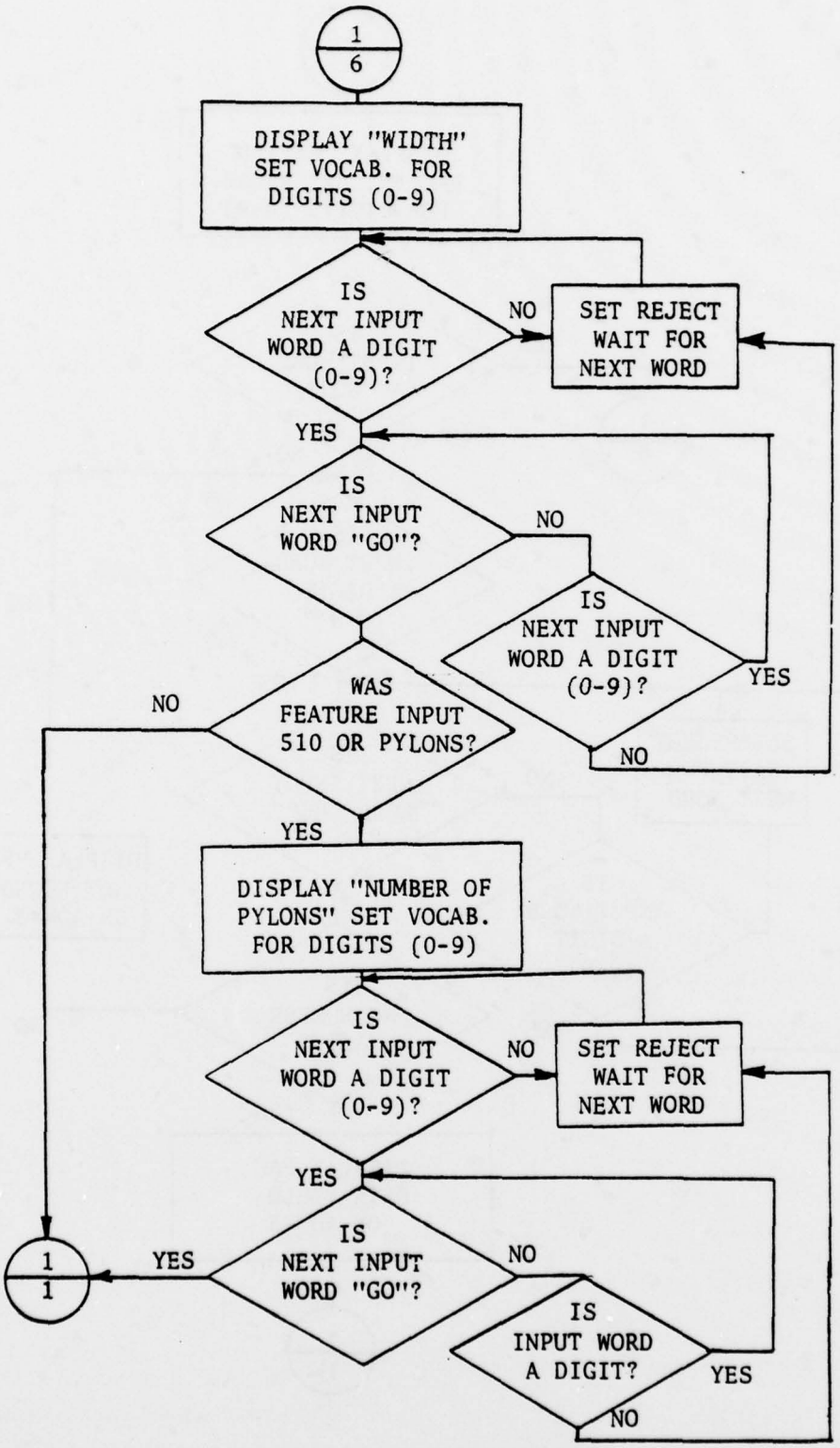


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 7 of 9)

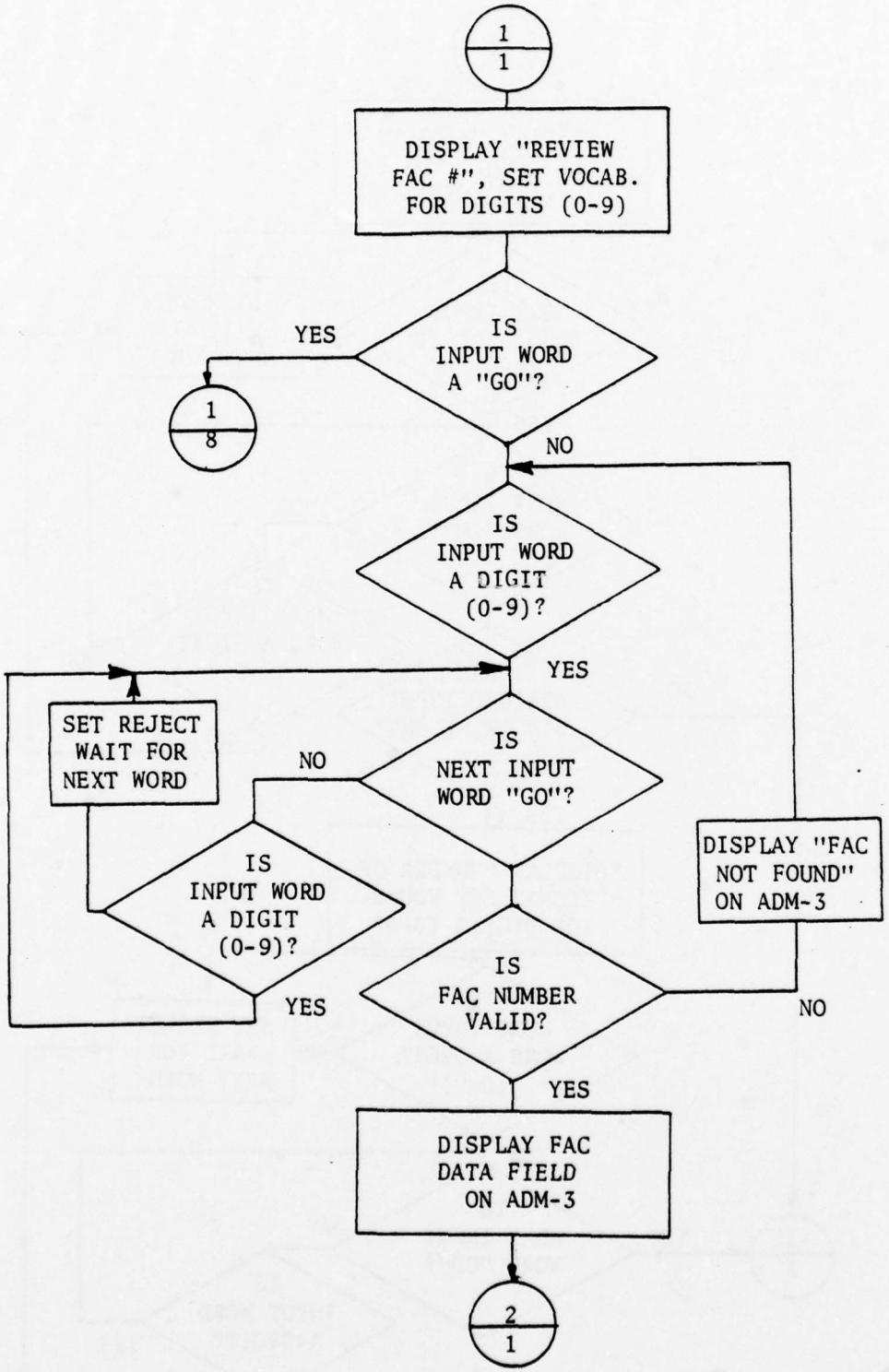


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 8 of 9)

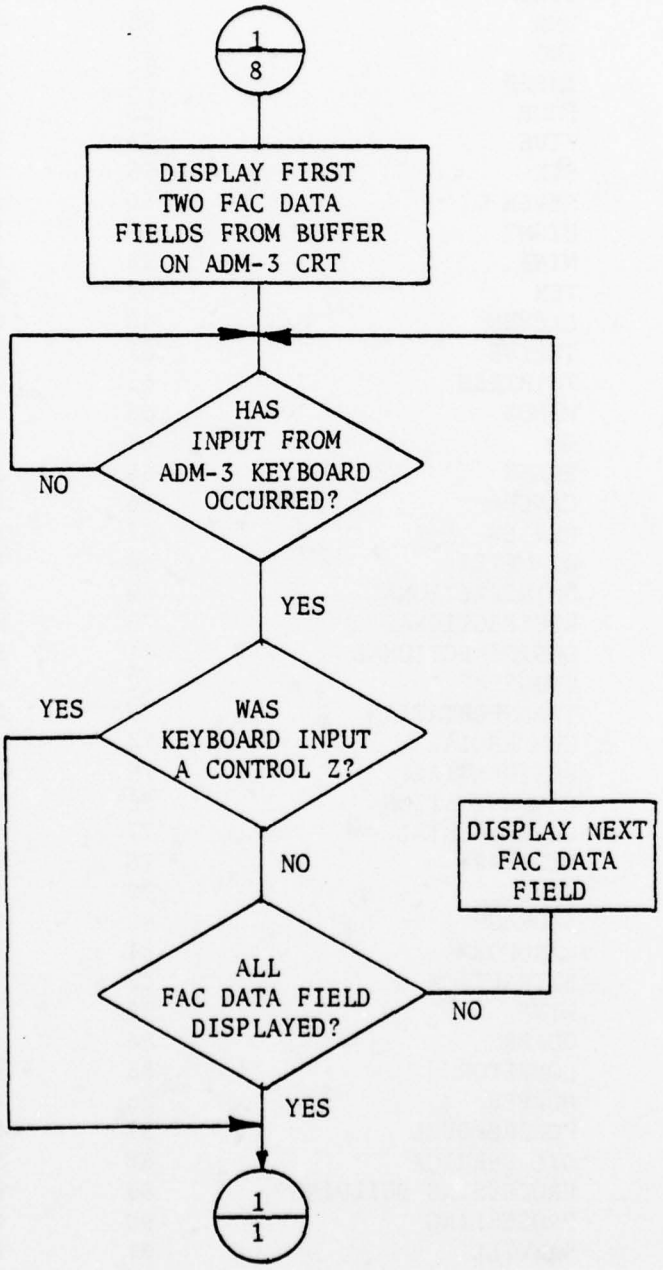


Fig. 2 Flowchart of DRLMS Fixed Node Program (Page 9 of 9)

TABLE I DRLMS WORD LIST (Sheet 1)

WORD NUMBER	WORD DISPLAYED	WORD NUMBER	WORD DISPLAYED
0	ZERO	49	STACK
1	ONE	50	ELECTRIC POWER
2	TWO	51	COOLING TOWER
3	THREE	52	SEWAGE TREATMENT
4	FOUR	53	PUMPING STATION
5	FIVE	54	FABRICATION
6	SIX	55	STEEL PLANT
7	SEVEN	56	SHIPBUILDING
8	EIGHT	57	LOCOMOTIVE MANUFACTURING
9	NINE	58	AIRCRAFT
10	TEN	59	MOTOR VEHICLE
11	ELEVEN	60	DISPOSAL
12	TWELVE	61	METAL ORE HEAP
13	THIRTEEN	62	COAL STORAGE
14	MINUS	63	SCRAP YARD
15	GO	64	TAILINGS
16	ERASE	65	RAILROAD YARD
17	CANCEL	66	RAIL SIDING
18	REVIEW	67	RAILROAD EMBANKMENT
19	DIAMETER	68	ELEVATED TRACK
20	UNIDIRECTIONAL	69	ELECTRIC RAILROAD
21	BIDIRECTIONAL	70	HIGHWAY
22	OMNIDIRECTIONAL	71	ELEVATED ROAD
23	INDUSTRY	72	CAUSEWAY
24	TRANSPORTATION	73	PIPELINE
25	COMMERCIAL	74	ABOVE GROUND
26	RESIDENTIAL	75	PENSTOCK
27	COMMUNICATION	76	BRIDGE
28	GOVERNMENTAL	77	ROAD BRIDGE
29	MILITARY	78	RAILROAD BRIDGE
30	CIVIL	79	OVERPASS
31	STORAGE	80	ROAD OVERPASS
32	LANDFORM	81	RAILROAD OVERPASS
33	EXTRACTION	82	BUSINESS
34	MINE	83	STORE
35	QUARRY	84	SHOPPING CENTER
36	CONVEYOR	85	BANK
37	HOPPER	86	OFFICE
38	POWERSHOVEL	87	RECREATION
39	OIL DERRICK	88	STADIUM
40	PROCESSING BUILDING	89	DRIVE-IN
41	PROCESSING	90	GRAND-STAND
42	SAWMILL	91	DISPLAY SIGN
43	COKE PLANT	92	MULTI-FAMILY
44	PETROLEUM	93	APARTMENT
45	REFINERY	94	SINGLE FAMILY
46	SYNTHETIC OIL	95	HOUSE
47	IRON SMELTER	96	MOBILE HOME
48	CHIMNEY	97	AGRICULTURE

TABLE I DRLMS WORD LIST (Sheet 2)

WORD NUMBER	WORD DISPLAYED	WORD NUMBER	WORD DISPLAYED
98	FARM HOUSE	138	PORT FACILITY
99	BARN	139	DAM
100	SHED	140	LOCK
101	POWER	141	LIGHT HOUSE
102	POWERLINE	142	GROUND FACILITY
103	TRANSFORMER	143	BARRACKS
104	SUBSTATION	144	MOTOR POOL
105	TRANSMITTING	145	GARAGE
106	RECEIVING	146	DEPOT
107	RADIO MAST	147	TANKS
108	RADIO TOWER	148	PETROLEUM TANK
109	RADAR SCREEN	149	WATER TANK
110	RADIO-NAV-AID	150	GAS HOLDER
111	UNI-RADAR	151	SILO
112	BI-RADAR	152	GRAIN ELEVATOR
113	OMNI-RADAR	153	STORAGE SILO
114	CITY HALL	154	DESERT
115	COURT BUILDING	155	ROCK
116	GOVERNMENT BUILDING	156	SAND
117	INSTITUTION	157	NORMAL SOIL
118	SCHOOL	158	MARSH
119	CHURCH	159	SWAMP
120	HOSPITAL	160	VEGETATION
121	AIRFIELD	161	MIXED WOODS
122	RUNWAY	162	ORCHARD
123	PARKING AREA	163	HEDGEROW
124	PARKING LOT	164	DECIDUOUS TREES
125	LEVEL CONCRETE	165	CONIFEROUS TREES
126	HANGAR	166	WATER
127	TERMINAL BUILDING	167	RIVER
128	CONTROL TOWER	168	LAKE
129	HEADQUARTERS BUILDING	169	CANAL
130	NAVAL	170	DRY LAKE
131	SEA	171	WALL
132	WHARF	172	FENCE
133	PIER	173	LEVEE
134	WAREHOUSE	174	EARTH EMBANKMENT
135	BREAKWATER	175	SNOW
136	JETTY	176	ICE
137	QUAY	177	CROPLAND

TABLE II FLIPS WORD LIST (Sheet 1)

WORD NUMBER	WORD	WORD NUMBER	WORD
0	0	41	OSCAR
1	1	42	PETER
2	2	43	QUEBEC
3	3	44	ROMEO
4	4	45	SIERRA
5	5	46	TANGO
6	6	47	UNIFORM
7	7	48	VICTOR
8	8	49	WHISKEY
9	9	50	X-RAY
10	NEXT	51	YANKEE
11	ERASE	52	ZULU
12	EXIT	53	VOR (V-O-R)
13	KEYBOARD	54	TACAN
14	10	55	VORTAC
15	EDIT MODE	56	NDB (N-D-B)
16	DELETE	57	MARKER BEACON
17	MODIFY	58	PROFILE-FAF
18	MOVE	59	RADIO RANGE
19	LINE	60	SIMUL-VOICE
20	WORD	61	NON-SIMUL
21	SYMBOL	62	REPORTING POINT
22	STATUS	63	COMPULSORY
23	CIRCLE	64	NON-COMPULSORY
24	ARC	65	MARKER IDENT
25	POINT ON LINE	66	TACAN-DME
26	AREA TYPE	67	INITIAL APPROACH
27	ALPHA	68	OBSTRUCTION
28	BRAVO	69	SPOT ELEVATION
29	CHARLIE	70	LEADERS
30	DELTA	71	PROCEDURE TURN BARB
31	ECHO	72	NON-DIRECTIONAL RADIO BEACON
32	FOXTROT	73	WARNING LIGHT
33	GEORGE	74	TEN-THOUSANDTHS
34	HOTEL	75	TWENTY-THOUSANDTHS
35	INDIA	76	LARGE
36	JULIE	77	MEDIUM
37	KILO	78	SMALL
38	LIMA	79	OPEN
39	MOTHER	80	CLOSED
40	NOVEMBER	81	HOLDING PATTERN

TABLE II FLIPS WORD LIST (Sheet 2)

WORD NUMBER	WORD	WORD NUMBER	WORD
82	WITH ARROWS	116	315 (three-fifteen)
83	WITHOUT ARROWS	117	SUBFILE LAYER
84	ALTERNATE MINIMUMS	118	LINE WEIGHT
85	TAKEOFF MINIMUMS	119	SMOOTHING
86	ARROW HEAD	120	ON
87	DISPLACE THRESHOLD	121	OFF
88	JET BARRIER	122	JUSTIFY TEXT
89	ARRESTING GEAR	123	ABOVE
90	CONTROL TOWER	124	BELOW
91	ROTATING BEACON	125	FILE CONTROL
92	VISUAL FLIGHT PATH	126	DRAWING
93	METAL SURFACE	127	SCRATCH
94	HELICOPTER LANDING	128	NEW
95	PROCEDURE TRACK	129	ACTIVATE AND DISPLAY
96	ILS (I-L-S)	130	CLEAR
97	LOCALIZER	131	STORE AND CLEAR
98	TRAP MOVE	132	RECALL
99	TRAP DELETE	133	LIST
100	ALL POINTS	134	PLACEMENT
101	OPERATION	135	NORMAL
102	SCALE SYMBOL	136	PLUS
103	GRID ROUNDOFF	137	MINUS
104	ROUNDNESS CONTROL	138	DECIMAL
105	GRAPHIC MODE	139	RETURN
106	ARC CLOCKWISE	140	SPACE
107	ARC C-C-W	141	BLANK
108	STREAM DIGITIZE	142	SOLID
109	ROTATE SYMBOL	143	HALF-TONE
110	45 (forty-five)	144	CLOCKWISE
111	90 (ninety)	145	C-C-W
112	135 (one-thirty-five)	146	TETRAHEADRON
113	180 (one-eighty)	147	LANDING TEE
114	225 (two-twenty-five)	148	ASTERISKS
115	270 (two-seventy)	149	APPROACH LIGHTING

TABLE III

ORGANIZATION OF FLIPS VOCABULARY INTO CATEGORIES (Sheet 1 of 4)

Category Name: FILE CONTROL

Valid Words:

- DRAWING (sub-category name)
  - Valid in sub-category:
    - NEW
    - ACTIVATE & DISPLAY
    - CLEAR
    - STORE & CLEAR
    - RECALL
    - LIST
- SCRATCH
  - ACTIVATE & DISPLAY
  - CLEAR
  - STORE & CLEAR
  - RECALL
  - LIST
- SYMBOL
  - NEW
  - NORMAL
  - RECALL
  - LIST

Category Name: EDIT MODE

Valid Words:

- DELETE
  - LINE
  - TEXT
  - SYMBOL
  - CIRCLE
  - ARC
- MODIFY
  - LINE
  - TEXT
  - SYMBOL
  - CIRCLE
  - ARC
  - STATUS
- MOVE
  - POINT ON LINE
  - TEXT
  - SYMBOL
  - ALL POINTS
  - TRAP MOVE
  - TRAP DELETE

TABLE III

ORGANIZATION OF FLIPS VOCABULARY INTO CATEGORIES (Sheet 2 of 4)

Category Name: KEYBOARD

Valid Words:
- ALPHA - ZULU (Phonetic Alphabet)
- 0-9 (Digits)
- PLUS
- MINUS
- DECIMAL
- SPACE
- RETURN

Category Name: OPERATION

Valid Words:
- SCALE SYMBOL
  - 0-9 (Digits)
  - DECIMAL
- GRID ROUNDOFF
  - 0-9 (Digits)
- GRAPHIC MODE
  - LINE
  - CIRCLE
  - ARC CW
  - ARC CCW
  - STREAM DIGITIZE
- ROTATE SYMBOL
  - 0
  - 45 (Forty-five)
  - 90 (Ninety)
  - 135 (etc.)
  - 180
  - 225
  - 270
  - 315
  - 0-9 (Digits)
- JUSTIFY TEXT
  - ABOVE
  - ON
  - BELOW
- SMOOTHING
  - ON
  - OFF
- SUBFILE LAYER
  - 0-9 (Digits)

TABLE III

ORGANIZATION OF FLIPS VOCABULARY INTO CATEGORIES (Sheet 3 of 4)

Category Name: LINE WEIGHT

Valid Words:
- 0-9 (Digits)
- AREA TYPE
  - BLANK
  - SOLID
  - HALFTONE

Category Name: SYMBOLS

Valid Words:
- VOR
- TACAN
- VORTAC
- NON-DIRECTIONAL BEACON
- MARKER BEACON
- RADIO RANGE
  - SIMUL-VOICE
  - NON-SIMUL
- REPORTING POINT
  - NON-COMPULSORY
  - COMPULSORY
- MARKER IDENT
- TACAN-DME
- INITIAL APPROACH
- OBSTRUCTION
- SPOT ELEVATION
- LEADERS
- PROCEDURE TURN
- CIRCLE
  - SMALL
  - MEDIUM
  - LARGE
  - OPEN
  - CLOSED
- HOLDING PATTERN
  - WITH ARROWS
  - NO ARROWS
  - CLOSED
  - OPEN
  - SMALL
  - MEDIUM
  - LARGE
  - CLOCKWISE
  - CCW

TABLE III

ORGANIZATION OF FLIPS VOCABULARY INTO CATEGORIES (Sheet 4 of 4)

Category Name: SYMBOLS (Continued)

Valid Words:

ALTERNATE MINIMUMS
TAKEOFF MINIMUMS
ARROW HEAD
DISPLAY THRESHOLD
JET BARRIER
ARRESTING GEAR
CONTROL TOWER
ROTATING BEACON
VISUAL FLIGHT PATHS
METAL SURFACE
HELICOPTER LANDING AREA
PROCEDURE TRACK
TETRAHEADRON
LANDING TEE
ASTERISKS
ILS
LOCALIZER
APPROACH LIGHTING
0-9 (Digits)
ALPHA-ZULU (Phonetic Alphabet)

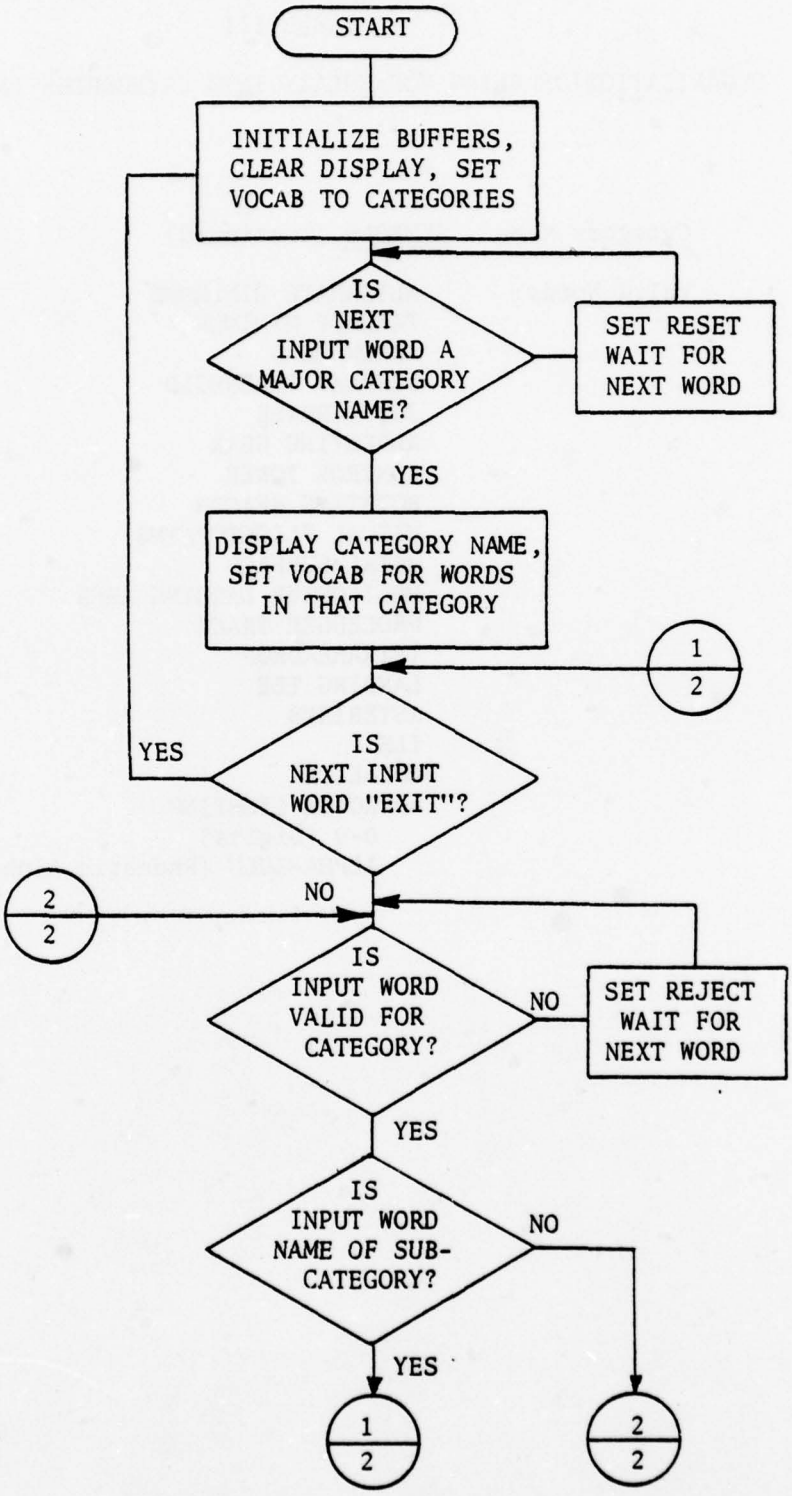


Fig. 3 FLIPS Program Flowchart (Page 1 of 2)

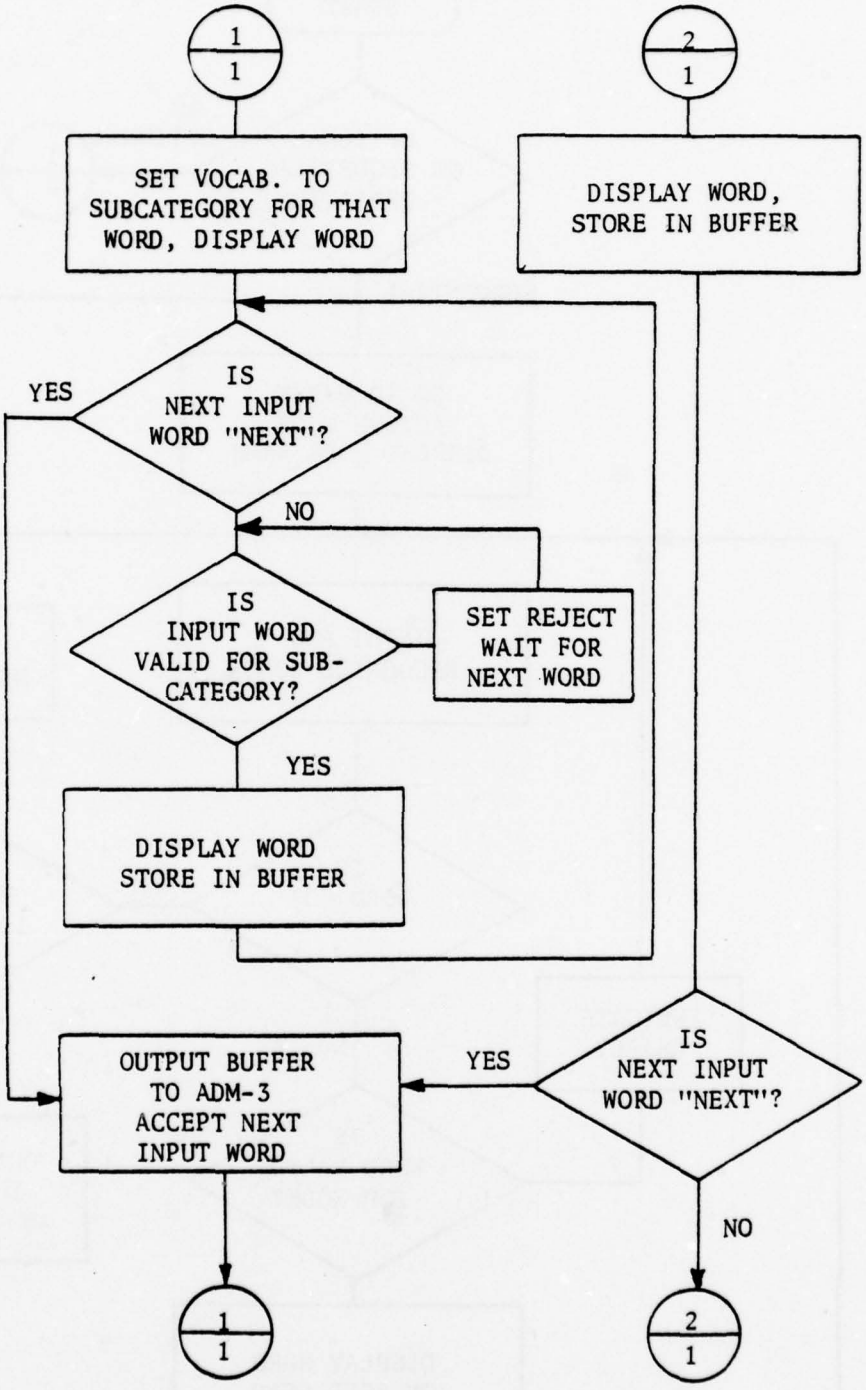


Fig. 3 FLIPS Program Flowchart (Page 2 of 2)

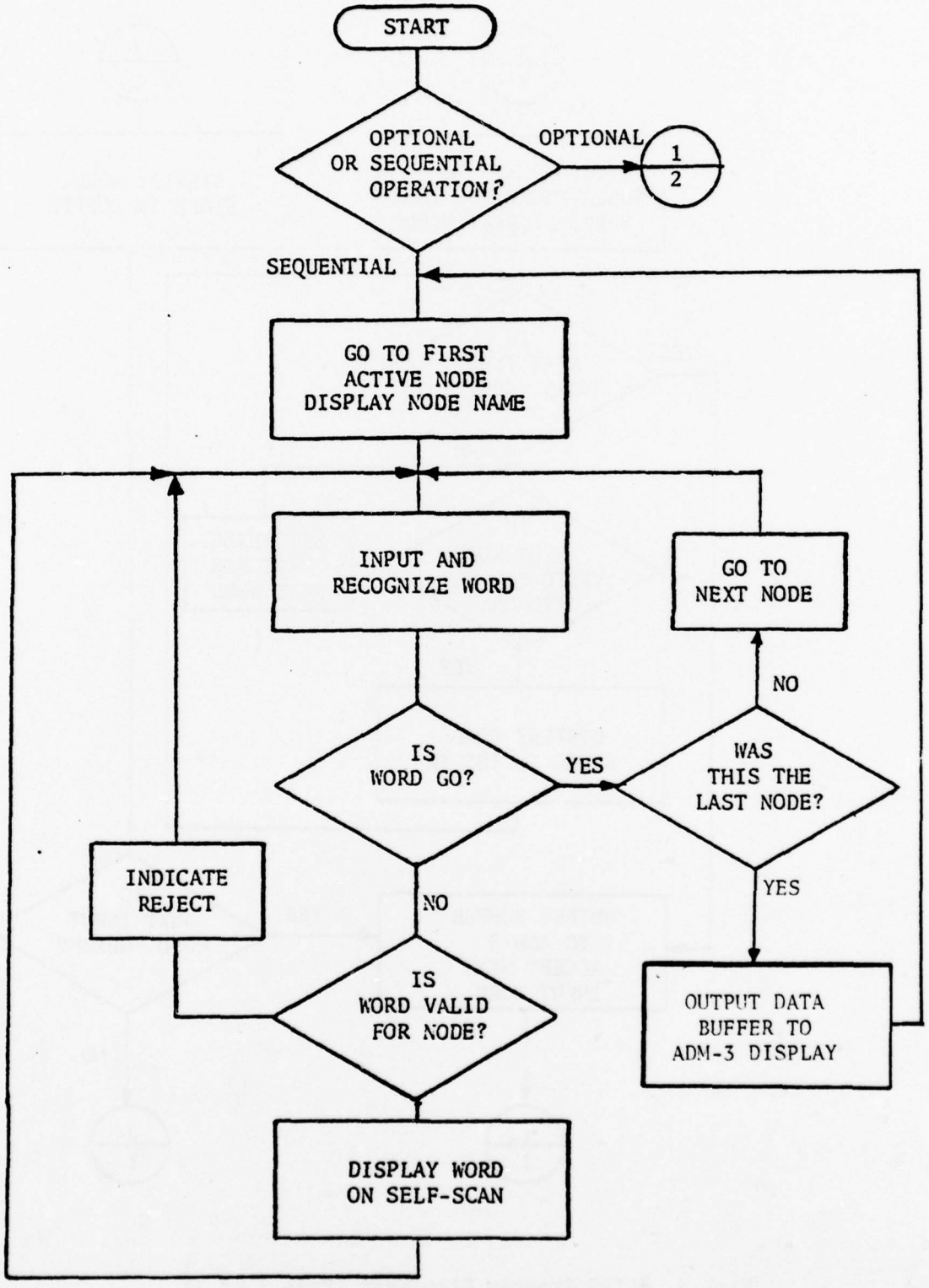


Fig. 4 Flowchart of DRLMS Adjustable Node Program (Page 1 of 3)

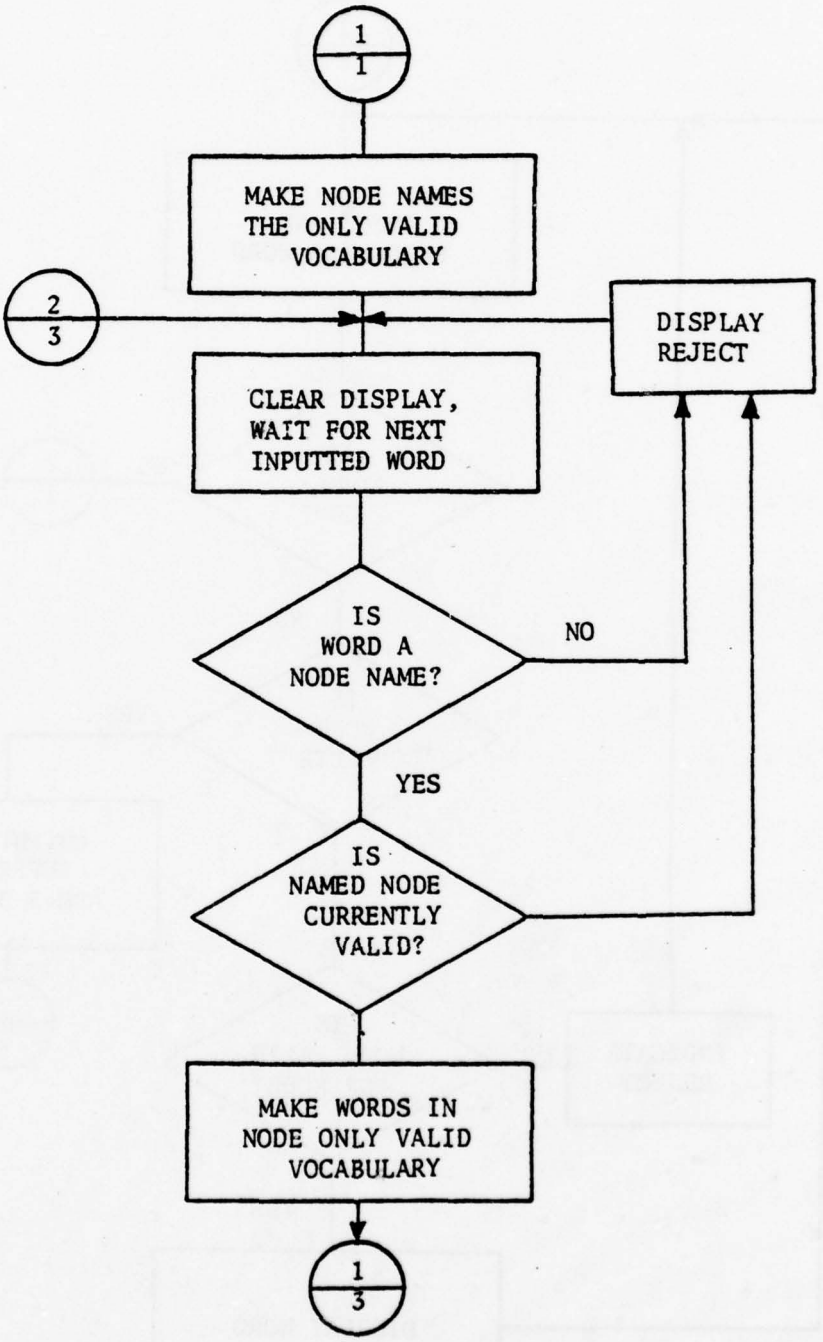


Fig. 4 Flowchart of DRLMS Adjustable Node Program (Page 2 of 3)

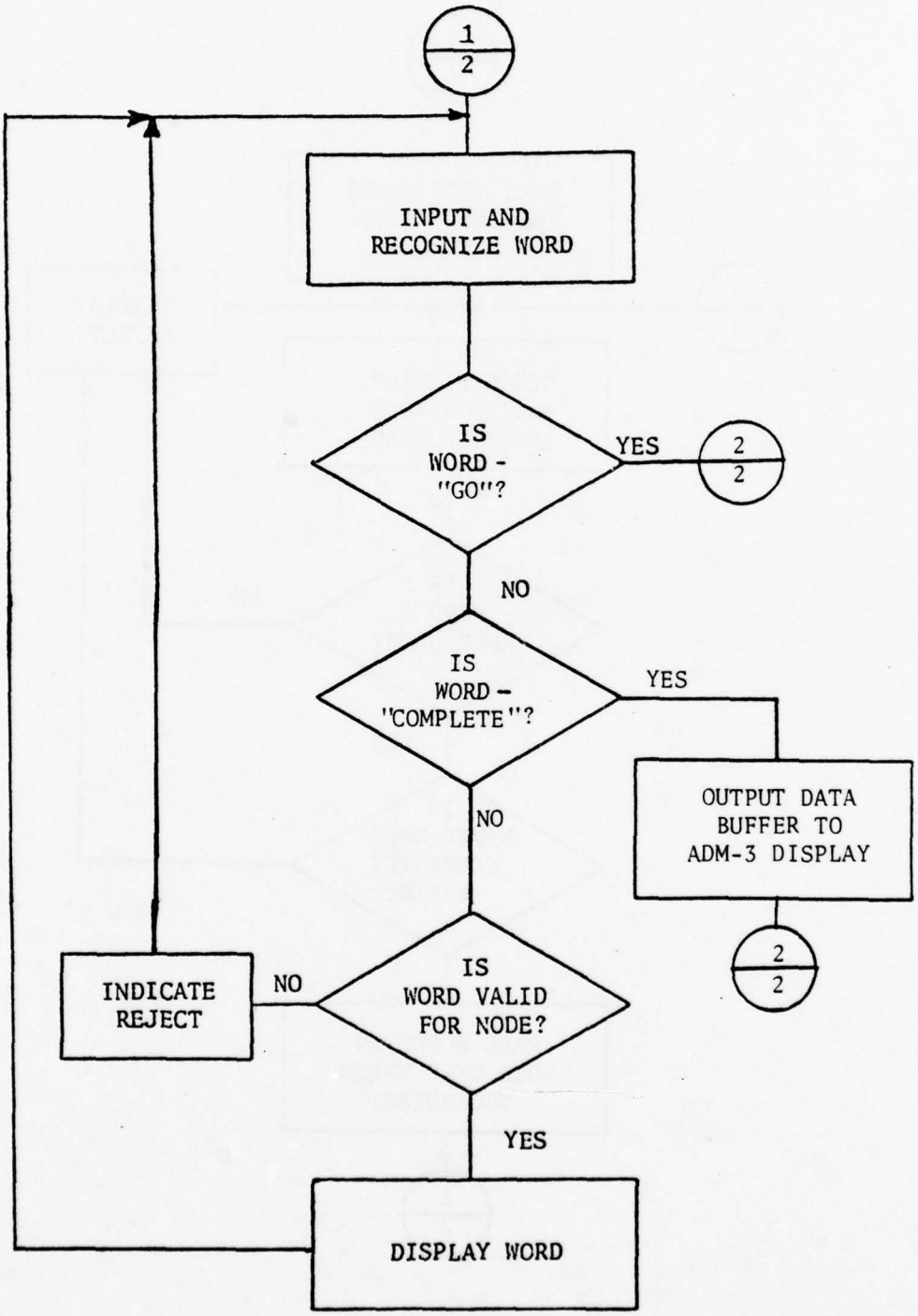


Fig. 4 Flowchart of DRLMS Adjustable Node Program (Page 3 of 3)

for vocabulary words are made through the Teletype input. Data fields are displayd on the ADM-3 CRT terminal screen. Up to and including 30 nodes may be included in the structure with a maximum of 100 words in each node. A particular vocabulary word may be included in any number of nodes. Every vocabulary word has a number from 0 through 499 which must be used for constructing nodes, for training and for constructing display messages. Any vocabulary word (except 0 through 13) may be represented on the output display by up to 16 letters (or digits or combination thereof). Vocabulary words 0 through 9 are the digits, word 10 is the command word GO, 11 is the command ERASE, word 12 is the command CANCEL and word 13 is the command COMPLETE. All other words are chosen by the operator. The command GO is always active. It serves to terminate the active node and to make active the next node in sequential or allow input of a new node name in optional operation. The word COMPLETE serves to display the previously inputted data field on the CRT (in the optional operation).

Once a particular node structure has been established, two types of operation are possible, Sequential and Optional. In the sequential operation, a talker must follow the predetermined sequence of nodes when inputting speech data. Any number of words in a particular node may be inputted. Each node is terminated and the next node in the sequence is made active by the command word GO. After the last node is terminated by a "GO", the completed data field is displayed automatically on the ADM-3 CRT display.

When the system is set into the sequential type of operation the name of the first active node will appear on the display. After each GO command is given, the name of the next node will appear on the display. The order of nodes in this operation is from lower to higher node numbers. When the program starts in the sequential mode, the lowest numbered node is always active first. After all nodes have been completed, the program reverts again to the lowest numbered node. A complete record of data inputted for the active nodes will then automatically appear on the ADM-3 CRT display.

The optional node type of operation allows the operator to choose by voice command, any of up to 30 nodes for use at a particular time. The node chosen is made active by inputting the name of the node. Another node may be subsequently made operative by exiting from the current node by the use of the GO command and then speaking the name of the next node desired. Only words 19 through 48 may be used to call nodes in optional operation. Node names for display can be the same as the vocabulary words used to call them. This is not manatory. The node names are assigned when constructing the node plan. Vocabulary word names are assigned independently. Node names appear on the display in optional operation when a node is called by the proper vocabulary word (from 19 to 48). Vocabulary word names appear on the display when inputting data or during training. Inputting data to a node can be terminated by a GO command. GO is always active in both types of operation. The command ERASE is always active and erases the last spoken word including node names in optional operation. In sequential operation, the node number and name cannot be erased. In sequential operation, CANCEL deletes all spoken inputs in a node but cannot delete the node.

At any time during the optional operation a complete record of inputted data may be displayed on the ADM-3 CRT provided that the current node is first terminated by a GO command. The command COMPLETE is then spoken. The display will indicate the message "COMPLETE?" to ensure that that function is desired by the operator. A GO command will then cause the display of the data on the CRT. This data will then be destroyed so that the next time COMPLETE is used only new data will be displayed.

Section III

SYSTEM ACCURACY TESTS

A. Background

Preliminary and final tests of the VRS-ADM were conducted at the TTI facility at Delran, NJ and at RADC. The preliminary tests, held at TTI, were designed to establish system accuracy on both vocabularies and were held shortly before the equipment was shipped to RADC. The final acceptance test was conducted by RADC personnel at RADC after the system was delivered and installed. Both test series involving the DRLMS vocabulary were conducted by the use of a special variation of the DRLMS fixed node program. This version allowed any number of input words to be spoken in any node of the program, provided such words were valid for that particular node. The command word "GO" was always valid and served to step the program from node to node when needed. The FLIPS program was not changed for the FLIPS test series which was conducted only at TTI. For each of the three test series, 10 talkers of both sexes were used.

The actual spoken inputs in the DRLMS tests both at TTI and at RADC were in accordance with the special DRLMS test vocabulary arrangement of Table IV. During each test session the test subject was instructed to read each word in this vocabulary list exactly in the order that it appeared. This list provided inputs of each and every word in each and every possible node of the DRLMS program except for two nodes. These two nodes are "Percent of Roof" and "Percent of Trees". These two nodes were deleted because a large number of digits only were required by other nodes. Certain vocabulary words occurred much more often than others in this list. For example, each feature identifier word occurs one time, the digits 0 through 9 seven times with certain digits appearing in other nodes. This arrangement simulates closely the actual vocabulary usage in an operational situation.

A somewhat different approach was used in the arrangement of the FLIPS test vocabulary shown in Table V. Except for the control words NEXT and EXIT care was taken to minimize repetition of words which were valid in more than one node of this program. This approach was taken by TTI who constructed both lists to provide a greater variety of test conditions.

Training procedures varied somewhat between the preliminary tests conducted at TTI and the acceptance tests conducted at RADC. At TTI, each person participating in the DRLMS test and seven of ten who participated in the FLIPS test inputted training data six to eight days before the test sessions were conducted. The training for these talkers consisted of five repetitions of each vocabulary word. The remaining three talkers in the FLIPS test inputted training data, two repetitions for each vocabulary word, and then tested the system immediately. Audio tape recordings of both training and test sessions were made for each talker. The 8035 Audio Input/Remote Control unit supplied with the VRS-ADM was modified to allow tape recordings to be made and played back through the system easily. These recordings were delivered with the system to RADC. Training reference patterns were also stored on digital magnetic

TABLE IV DRLMS TEST VOCABULARY (sheet 1)

DISPLAY PROMPTS	TEST WORDS	
FAC. NUMBER	0 through 9	
	GO	
IDENTIFY FEATURE	INDUSTRY	ELEVATED TRACK
	TRANSPORTATION	ELECTRIC RAILROAD
	COMMERCIAL	HIGHWAY
	RESIDENTIAL	ELEVATED ROAD
	COMMUNICATION	CAUSEWAY
	GOVERNMENTAL	PIPELINE
	MILITARY	ABOVE GROUND
	CIVIL	PENSTOCK
	STORAGE	BRIDGE
	LANDFORM	ROAD BRIDGE
	EXTRACTION	RAILROAD BRIDGE
	MINE	OVERPASS
	QUARRY	ROAD OVERPASS
	CONVEYOR	RAILROAD OVERPASS
	HOPPER	BUSINESS
	POWERSHOVEL	STORE
	OIL DERRICK	SHOPPING CENTER
	PROCESSING BUILDING	BANK
	PROCESSING	OFFICE
	SAWMILL	RECREATION
	COKE PLANT	STADIUM
	PETROLEUM	DRIVE-IN
	REFINERY	GRAND-STAND
	SYNTHETIC OIL	DISPLAY SIGN
	IRON SMELTER	MULTI-FAMILY
	CHIMNEY	APARTMENT
	STACK	SINGLE FAMILY
	ELECTRIC POWER	HOUSE
	COOLING TOWER	MOBILE HOME
	SEWAGE TREATMENT	AGRICULTURE
	PUMPING STATION	FARM HOUSE
	FABRICATION	BARN
	STEEL PLANT	SHED
	SHIPBUILDING	POWER
	LOCOMOTIVE MANUFACTURING	POWERLINE
	AIRCRAFT	TRANSFORMER
	MOTOR VEHICLE	SUBSTATION
	DISPOSAL	TRANSMITTING
	METAL ORE HEAP	RECEIVING
	COAL STORAGE	RADIO MAST
	SCRAP YARD	RADIO TOWER
	TAILINGS	RADAR SCREEN
	RAILROAD YARD	RADIO NAV AID
	RAIL SIDING	UNI-RADAR
	RAILROAD EMBANKMENT	BI-RADAR

TABLE IV DRLMS TEST VOCABULARY (sheet 2)

DISPLAY PROMPTS	TEST WORDS
IDENTIFY FEATURE (CONT'D)	OMNI RADAR
	CITY HALL
	COURT BUILDING
	GOVERNMENT BUILDING
	INSTITUTION
	SCHOOL
	CHURCH
	HOSPITAL
	AIRFIELD
	RUNWAY
	PARKING AREA
	PARKING LOT
	LEVEL CONCRETE
	HANGAR
	TERMINAL BUILDING
	CONTROL TOWER
	HEADQUARTERS BUILDING
	NAVAL
	SEA
	WHARF
	PIER
	WAREHOUSE
	BREAKWATER
	JETTY
	QUAY
	PORT FACILITY
	DAM
	LOCK
	LIGHT HOUSE
	GROUND FACILITY
	BARRACKS
	MOTOR POOL
	GARAGE
	DEPOT
	TANKS
	PETROLEUM TANK
	WATER TANK
	GAS HOLDER
	SILO
	GRAIN ELEVATOR
	STORAGE SILO
	DESERT
	ROCK
	SAND
	NORMAL SOIL
	MARSH
	SWAMP
	VEGETATION
	MIXED WOODS
	ORCHARD
	HEDGEROW
	DECIDUOUS TREES
	CONIFEROUS TREES
	WATER
	RIVER
	LAKE
	CANAL
	DRY LAKE
	WALL
	FENCE
	LEVEE
	EARTH EMBANKMENT
	SNOW
	ICE
	CROPLAND
	GO

TABLE IV DRLMS TEST VOCABULARY (sheet 3)

DISPLAY PROMPTS	TEST WORDS
SURFACE MATERIAL	1 through 13 GO
PREDOM. HEIGHT	0 through 9 GO
FEATURE TYPE	0 1 2 GO
ORIENTATION	0 through 9 GO
LENGTH OR DIA.	0 through 9 GO
WIDTH	0 through 9 GO
NUMBER OF PYLONS	0 through 9 GO
DIRECTIVITY	UNIDIRECTIONAL BIDIRECTIONAL OMNIDIRECTIONAL GO
WIDTH	0 through 9 GO
BUILDING DENSITY	1 through 13 CANCEL

TABLE V FLIPS TEST VOCABULARY (Sheet 1)

FILE CONTROL	OPERATION	KEYBOARD
DRAWING	SCALE SYMBOL	PLUS
NEW	NEXT	MINUS
ACTIVATE AND DISPLAY	GRID ROUNDOFF	SPACE
CLEAR	NEXT	DECIMAL
STORE AND CLEAR	GRAPHIC MODE	RETURN
RECALL	LINE	0 through 9
NEXT	CIRCLE	ALPHA
SCRATCH	ARC CLOCKWISE	BRAVO
LIST	ARC C-C-W	CHARLIE
NEXT	STREAM DIGITIZE	DELTA
SYMBOL	NEXT	ECHO
NEXT	ROTATE SYMBOL	FOXTROT
EXIT	45 (forty-five)	GEORGE
	90 (ninety)	HOTEL
EDIT MODE	135 (one-thirty-five)	INDIA
	180 (one-eighty)	JULIE
DELETE	225 (two-twenty-five)	KILO
NEXT	270 (two-seventy)	LIMA
MODIFY	315 (three-fifteen)	MOTHER
LINE	NEXT	NOVEMBER
SYMBOL	JUSTIFY TEXT	OSCAR
CIRCLE	ABOVE	PETER
ARC	ON	QUEBEC
STATUS	BELOW	ROMEO
NEXT	NEXT	SIERRA
MOVE	SMOOTHING	TANGO
POINT ON LINE	NEXT	UNIFORM
ALL POINTS	SUBFILE LAYER	VICTOR
TRAP MOVE	NEXT	WHISKEY
TRAP DELETE	EXIT	X-RAY
WORD		YANKEE
NEXT		ZULU
EXIT		NEXT
		EXIT
LINE WEIGHT		
AREA TYPE		
BLANK		
SOLID		
HALF TONE		
NEXT		
EXIT		

TABLE V FLIPS TEST VOCABULARY (Sheet 2)

SYMBOL

RADIO RANGE	HOLDING PATTERN
SIMUL-VOICE	WITH ARROWS
NON-SIMUL	WITHOUT ARROWS
NEXT	CLOSED
REPORTING POINT	OPEN
NON-COMPULSORY	SMALL
COMPULSORY	MEDIUM
NEXT	LARGE
MARKER IDENT	CLOCKWISE
TACAN-DME	C-C-W
OBSTRUCTION	NEXT
SPOT ELEVATION	METAL SURFACE
NEXT	HELICOPTER LANDING
CIRCLE	PROCEDURE TRACK
SMALL	TETRAHEADRON
MEDIUM	LANDING TEE
LARGE	ASTERISKS
OPEN	I-L-S
CLOSED	LOCALIZER
NEXT	LEADERS
ALTERNATE MINIMUMS	PROCEDURE TURN BARB
ROTATING BEACON	V-O-R
TAKEOFF MINIMUMS	TACAN
ARROW-HEAD	VORTAC
DISPLACE THRESHOLD	MARKER BEACON
JET BARRIER	N-D-B
ARRESTING GEAR	APPROACH LIGHTING
CONTROL TOWER	NEXT
NEXT	EXIT

tape for each talker for each test. This tape was also supplied to RADC with the system.

Training for the RADC tests was accomplished a day before testing. Each talker spoke each vocabulary word 10 times into the VRS-ADM. The resulting reference array patterns were stored on digital magnetic tape.

B. Test Results

Similar procedures were used for the actual test sessions held at both TTI and RADC. Reference data for a talker to be tested was read from digital magnetic tape into the memory of the Nova 800. The talker was then asked to read the test vocabulary list of the vocabulary for which he or she had previously established reference data. The DRLMS list included 267 words, the FLIPS list had 173 words. Each talker was instructed that if any word spoken was misrecognized or rejected, to repeat the word once. If the word was misrecognized or rejected again, the talker was instructed to proceed to the next word. After completion of the first pass through the list, the talker retrained with five repetitions any words which were misrecognized or rejected. The talker was then instructed to repeat the test list words following the same procedures as the first pass.

Retraining after the first pass was done in order to simulate field operation of a new word recognition system. It has been the experience of TTI that new operators or even experienced operators need a short period of acclimation to a new vocabulary. This acclimation includes retraining of a few words which have been found to cause recognition problems.

The person conducting the test made note of all misrecognitions and rejects as well as corrections which could be accomplished by repeating one time a misrecognized or rejected word. Misrecognition errors occurred when one word was spoken and another was recognized by the VRS-ADM. Rejects occurred when the system was unable to exceed a correlation threshold for any node vocabulary word or when a word outside a node was recognized.

1. Preliminary Tests at TTI

These tests were conducted by TTI to satisfy the accuracy requirements of the program. Eight male and two female talkers participated in the DRLMS vocabulary tests, with nine male and one female talker taking part in the FLIPS vocabulary test. Results for these two tests are tabulated in Table VI and VII for DRLMS and FLIPS respectively. These tables illustrate raw scores and corrected scores for each part of the tests and for the complete tests. Raw scores were calculated counting all rejects and misrecognition errors. The corrected scores were calculated by counting only rejects and misrecognition errors remaining after the words causing these errors were repeated one time. The design goal of 99% recognition accuracy was very nearly reached (within 1%) by the combined raw scores of the 10 talkers in the DRLMS test and exceeded by the combined corrected scores of the talker in both tests.

In the DRLMS test only one talker (#10) had mediocre results. This

TABLE VI RESULTS OF DRLMS TEST AT TTI

	First Repetition					Second Repetition					Complete Test	
Spkr.	Errors	Rejects	Errors & Rejects Corrected	Raw Score (%)	Corrected Score (%)	Errors	Rejects	Errors & Rejects Corrected	Raw Score (%)	Corrected Score (%)	Raw Score (%)	Corrected Score (%)
1	3	0	0	98.9	98.9	0	1	0	99.6	99.6	99.2	99.2
2	3	1	1	98.5	98.9	1	0	1	99.6	100.0	99.1	99.4
3	2	0	1	99.3	99.6	2	0	0	99.3	99.3	99.3	99.4
4	3	2	2	98.1	98.9	2	0	0	99.3	99.3	98.7	97.1
5	6	0	4	97.8	99.3	5	0	5	98.1	98.1	97.9	96.6
6	4	0	0	98.5	98.5	1	0	0	99.6	96.6	99.1	99.1
7	2	0	1	99.3	99.6	7	1	7	97.0	99.6	98.1	99.6
8*	6	1	3	97.4	98.5	5	0	2	98.1	98.9	97.8	98.7
9	5	0	3	98.1	99.3	1	0	1	99.6	100.0	98.9	99.6
10*	15	0	13	94.4	95.2	10	0	8	96.3	99.3	95.3	99.2
Total	49	4	28	98.0	99.1	34	2	24	98.7	99.6	98.3	99.2

*Female talker

TABLE VII RESULTS OF FLIPS TEST AT TTI

	First Repetition					Second Repetition					Complete Test	
Spkr.	Errors	Rejects	Errors & Rejects Corrected	Raw Score (%)	Corrected Score (%)	Errors	Rejects	Errors & Rejects Corrected	Raw Score (%)	Corrected Score (%)	Raw Score (%)	Corrected Score (%)
1	0	2	0	98.8	98.8	0	3	2	98.3	99.4	98.6	99.1
2	2	4	5	96.5	99.4	5	4	9	94.8	100.0	95.7	99.7
3	0	0	0	100.0	100.0	0	0	0	100.0	100.0	100.0	100.0
4	2	0	1	98.8	99.4	1	1	2	98.8	100.0	98.8	99.7
5	1	6	6	96.0	99.4	1	3	3	97.7	99.4	96.8	99.4
6	5	4	6	94.8	98.3	3	0	2	98.3	99.4	96.5	98.8
7	2	3	1	97.1	97.7	0	4	3	97.7	99.4	97.4	98.6
8*	1	8	2	94.8	96.0	2	0	1	98.8	99.4	96.8	97.7
9	1	1	2	98.8	100.0	3	0	2	98.3	99.4	98.6	99.7
10	1	1	2	98.8	100.0	0	1	1	99.4	100.0	99.1	100.0
Total	15	29	25	97.5	98.9	15	16	25	98.2	99.7	97.8	99.3

*Female talker

talker, a female TTI employee has no appreciable experience inputting speech data to word recognition systems. She has, however, recorded test data on a few previous occassions. Of the other nine talkers in this test, five, numbers 1, 2, 3, 4 and 9, were very experienced in talking to word recognition systems, and the remaining four had some occassional experience in this area. All talkers were TTI employees.

The overall accuracies realized in the FLIPS tests were close to those of the TTI DRLMS test. Individual scores varied somewhat more in FLIPS than in DRLMS. One talker in the FLIPS test had no errors. He is an experienced speech researcher. All talkers in the test series except numbers 7 and 8 have had considerable experience talking to word recognition systems. Talker number 8, the only female and also the only non-employee in this test had never talked to a WRS before. In view of this fact and also the two repetition train which she performed just before testing, her accuracy was quite good. The variation in the training procedure did not appear to have any significant effect on recognition accuracy. Talkers 7 through 9 trained with two repetitions and tested immediately afterwards.

2. Analysis of Errors in DRLMS Test at TTI

In the 10 talker DRLMS test conducted at TTI, a total of 89 misrecognitions and rejects occurred before corrections. These errors involved a total of 51 different words of the DRLMS vocabulary of 178 words. Therefore, none of the remaining 127 vocabulary words were misrecognized for any of the 10 talkers. A complete list of the 51 words which were misrecognized or rejected is shown in Table VIII. The more traditional error matrix usually used for this type of illustration has been omitted because of the small portion of the total vocabulary in error. Also a 178 by 178 matrix would be rather large. The table is divided into three sections according to the vocabulary structure.

The first field in Table VIII is the "FAC Number" field in which only digits are valid for recognition. Only one error occurred in this field for the complete talker set, a substitution of ONE for a spoken FIVE. Each word in this field was spoken a total of 20 times in the test. The greatest number of errors occurred in the IDENTIFY FEATURE field. All except four of the vocabulary words are valid in this field. The four not valid are DIAMETER, UNDIRECTIONAL, BIDIRECTIONAL and OMNIDIRECTIONAL. The latter three of these four words are valid under the DIRECTIVITY field. Each word in the IDENTIFY FEATURE field was spoken two times by each talker for a total of 20 times in the test.

The errors of the nine remaining fields used in the DRLMS test (as shown in sheet 3 of table IV) have been listed together because of the similarity of these fields. Each of these nine fields includes some or all of the digits 0 through 13. The total number of times each word in these fields was spoken is shown on Sheet 2 of Table VIII. These numbers of times spoken vary from word to word because of the variations in the composition of the fields. Sheet 3 of Table IV illustrates the construction of these fields.

The center column of Table VIII indicates words of the various fields

TABLE VIII LIST OF MISRECOGNIZED WORDS
IN DRLMS TEST AT TTI (Sheet 1)

(Each word in the two fields shown on this sheet was spoken twice by each of the 10 talkers for a total of 20 repetitions in the test)

FIELD	WORD SPOKEN	WORD(S) RECOGNIZED
FAC Number	Five	One
Identify Feature	Governmental	Commercial (3)
	Extraction	Scrap Yard
	Mine	Nine
	Oil Derrick	*
	Processing Building	Eleveted Road
	Coke Plant	Crop Land
	Sewage Treatment	Steel Plant
	Pumping Station	Fabrication
	Fabrication	Vegetation, Recreation
	Steel Plant	Coke Plant
	Shipbuilding	Court Building
	Metal Ore Heap	Dry Lake, Nine
	Road Overpass	Railroad Overpass
	Bank	Marsh
	Drive-In	Dry Lake
	Apartment	* (2)
	Mobile Home	Wall
	Substation	Vegetation
	Uni-Radar	Omni-Radar
	Government Building	Governmental
	Institution	*
	School	Store
	Hospital	Normal Soil
	Sea	Quay (2)
	Warf	Four (4)
	Quay	Sea
	Lock	Rock
	Barracks	Church
	Tanks	Bank (2)
	Desert	Jetty
	Rock	Lock (7)
	Sand	Shed, Dam
	Swamp	Snow, Rock
	Deciduous Trees	Coniferous Trees
	Water	Quarry
	Lake	Water, Levee, R
	Canal	Erase
	Dry Lake	Lake, Drive-In
	Cropland	Powerline
	Dam	R

TABLE VIII LIST OF MISRECOGNIZED WORDS
IN DRLMS TEST AT TTI (Sheet 2)

FIELD	WORD SPOKEN	TOTAL REPS	WORD(S) RECOGNIZED
(All Other Fields)	One	180	R
	Three	160	Eight (2)
	Four	160	One
	Five	160	Nine (9), One (2), R
	Seven	160	Nine, R
	Eight	160	Three (2)
	Ten	40	Seven
	Thirteen	40	R
Directivity	Bidirectional	20	Omnidirectional (2)
	Omnidirectional	20	Bidirectional, Unidirectional
	Go	200	Zero (2), R

Notes: R indicates a reject, * indicates that a dropout occurred, numbers in parenthesis after a misrecognized word indicates number of times misrecognition occurred.

which were spoken, but for which misrecognitions or rejects occurred. The righthand column indicates the misrecognitions and rejects which occurred. Rejects are indicated by "R". In a few instances, errors occurred because of long pauses within words. Pauses, associated with stop consonants, which are greater than 150 milliseconds in length can cause word boundary circuitry in the speech processor to indicate the end of a word at the wrong time. The resulting "dropout" error has been indicated by an asterisk (*) in Table VIII. Only four of the 87 errors were dropout errors.

Many of the errors in this test series could be expected. Substitution of ROCK for LOCK is predictable. Use of the plural for ROCK would decrease this substitution tendency. Five-Nine and Three-Eight confusions are also fairly common. Expected confusions such as SEA and QUAY (pronounced "key") and MINE and NINE occurred only once each. Such confusions as BANK and MARSH, DESERT and JETTY, and METAL ORE HEAP and NINE must be ascribed to sloppy pronounciation, breath noises or other talker problems.

3. Analysis of Errors in FLIPS Test at TTI

The FLIPS vocabulary of 150 words was highly structured into six categories or nodes with subcategories under five of the six categories. Therefore, the table of misrecognized and rejected words for the FLIPS test, shown in Table IX, is somewhat more complicated than the equivalent table for the DRLMS test. For each misrecognized or rejected word both the major category and the subcategory for the word in error is shown. In the instances where a word in error is not a member of a subcategory, a dash appears in the column SUBCATEGORY. For example, under the category EDIT MODE, the word DELETE was misrecognized. This word in this instance was the name of a subcategory and therefore not viable within a subcategory. The word SYMBOL appears several times in the structure of the FLIPS vocabulary as shown in Table III. SYMBOL is the name of a major category as well as the name of a subcategory under the major category, FILE CONTROL. It also appears under two subcategories, DELETE and MODIFY, under the category, EDIT MODE. However, this word was missed only once in the FLIPS test. As shown in Table IX, SYMBOL was misrecognized once as CIRCLE when spoken in the subcategory, MODIFY under the category, EDIT MODE. Each word was spoken two times in each subcategory (or category) by each of the 10 talkers for a total of 20 repetitions in the complete test. The control word NEXT was spoken 44 times by each talker or 440 total repetitions for the test. The control word EXIT was spoken 12 times by each talker, but was never misrecognized. The digits, TWO and SIX were misrecognized or rejected in different categories and are shown as so in Table IX.

Of the total of 75 errors detected in the FLIPS test, 49 were rejects and 26 were misrecognitions. The large number of rejects was due to the method of structuring the vocabulary in software. In the FLIPS program, the entire vocabulary was used for recognition regardless of current category or subcategory. That is, the feature array for an input word was compared with the reference arrays for all vocabulary words. After a recognition was made, the recognized word was checked for validity in the currently active category or subcategory. If the word was valid, it was displayed. If it was not valid a reject was signalled. Rejects were also signalled if all correlation scores

TABLE IX LIST OF MISRECOGNIZED WORDS IN FLIPS TEST AT TTI

CATEGORY	SUBCATEGORY	SPOKEN WORD	MISRECOGNITION
Edit Mode	-	Delete	R
	Modify	Symbol	Circle
	Modify	Arc	R (2)
	-	Move	R
	Move	Trap Move	Trap Delete
	Move	Trap Delete	Trap Move
Line Weight	Area Type	Blank	R (2)
	-	Two	R
	-	Six	R
Operation	-	Scale Symbol	R
	Scale Symbol	Two	R
	Graphic Mode	Line	R (3)
	Rotate Symbol	Forty-Five	One-Thirty-Five
	Rotate Symbol	Ninety	Nine
	Rotate Symbol	Two-Twenty-Five	One-Thirty-Five
	Rotate Symbol	Two-Seventy	R
	-	Smoothing	R
	Smoothing	Off	On
Keyboard	(none under keyboard)	Space	Six
		Plus	R
		Two	R (2)
		Four	R
		Five	One, R (2)
		Six	Space
		Eight	R
		Alpha	R
		Delta	Alpha, Bravo, Mike
		Echo	R
		India	R
		Kilo	Zero
		Quebec	Peter
		Whiskey	Uncle, R (4)
		Yankee	Eight, R (2)
		Zulu	Two
Symbol	-	Reporting Point	Tacan-DME, R
	Circle	20 Thousands	Ten-Thousands
	-	Holding Pattern	*, R
	Holding Pattern	Without Arrows	With Arrows, R
	Holding Pattern	Medium	*, R
	-	Localizer	R
	-	Leaders	R
	-	Procedure Turn Barb	Procedure Track
	-	V-O-R	*
	-	Approach Lighting	Spot Elevation
(Open Set)		Lineweight	R (3)
(Open Set)		Keyboard	R
(Control Word)		Next	R (7)

were below the normal reject threshold. In both DRLMS programs, the feature array for the input word is compared with reference arrays for only those words which are valid under the current category or mode. The test results illustrate the advantages of each approach. With the FLIPS approach of wide open initial recognition, rejects outnumber misrecognitions. Restricting the reference arrays gives fewer rejects but more misrecognitions. This latter observation is especially true if an operator accidently speaks words included in the vocabulary but not valid in an active mode. The resulting forced choice will often but not always result in a misrecognition.

4. Final Tests at RADC

The test conducted at RADC by RADC personnel involved only the DRLMS vocabulary as spoken by 10 talkers, all Air Force employees. Two of these 10 were females, five (1, 2, 4, 5 and 7) were experienced with word recognition systems, the remainder were not. The overall accuracy achieved by the RADC talkers as shown in Table X was about four percentage points lower than that achieved in the DRLMS test at TTI. The level of experience with word recognition systems of the talkers involved must be assumed as having an important effort on test scores. Eight of the 10 talkers in the TTI test were experienced. Five of those eight could be considered "very experienced". Three of the experienced RADC talkers registered corrected accuracies of approximately 98%. However, the other two experienced RADC talkers had lower accuracies than the poorest TTI talker. Error analysis of the RADC tests was not possible because substitution error data were not made available for these tests.

TABLE X RESULTS OF DRLMS TEST AT RADC

	First Repetition					Second Repetition					Complete Test	
Spkr.	Errors	Rejects	Errors & Rejects Corrected	Raw Score (%)	Corrected Score (%)	Errors	Rejects	Errors & Rejects Corrected	Raw Score (%)	Corrected Score (%)	Raw Score (%)	Corrected Score (%)
1	6	0	0	97.8	97.8	5	0	4	98.1	99.6	98.0	98.7
2	9	0	2	96.7	97.4	6	0	0	97.8	97.8	97.2	97.6
3*	23	2	9	90.7	94.0	19	0	6	92.9	95.2	91.8	94.6
4	12	0	7	95.5	98.1	7	0	2	97.4	98.1	96.5	98.1
5	32	0	6	88.1	90.3	10	3	8	95.2	98.1	91.6	94.2
6	21	1	0	91.8	91.8	23	1	0	91.0	91.0	91.4	91.4
7	17	0	0	93.7	93.7	10	0	0	96.3	96.3	95.0	95.0
8	8	0	6	97.0	99.3	9	0	7	96.7	99.3	96.8	99.3
9	30	9	7	85.5	88.1	25	1	0	90.3	90.3	87.9	89.2
10*	8	0	4	97.0	98.5	3	1	2	98.5	99.3	97.8	98.9
Total	166	12	41	93.4	94.9	117	6	29	95.4	96.5	94.4	95.7

*Female talker

Section IV

CONCLUSIONS AND RECOMMENDATIONS

The major objectives of the Word Recognition program were achieved. The advanced development model word recognition system developed for cartographic use during this program has been demonstrated at a recognition accuracy better than 99% by experienced talkers. Three different structured recognition programs will provide RADC with the capability of demonstrating the use of speech recognition in a variety of cartographic applications. Such applications include inputting of DRLMS data and preparations of FLIPS charts. The combined system of hardware and software delivered by TTI to RADC provides the capability of recognitions of up to 500 words in syntactic structures. The system includes the capability for remote training by any speaker of any word or a complete vocabulary of words for a cartographic data input task. Response time of the system to any word in any vocabulary is less than 0.5 seconds. This fast response to large vocabulary sets is due partly to the use of a TTI hardware high speed correlator which is an integral part of the system hardware. Provision for storing large numbers of complete data fields for the DRLMS applications has been made. An operator can recall such stored data for review and modification when necessary.

This VRS-ADM, while a very useful tool for the study of speech application in cartography, cannot be used for actual field operations. No type of interfacing with cartographic data base computers was required or supplied for this program. In order that the effectiveness of voice data entry in cartography be fully evaluated a WRS suitable for field operation should be developed. Such a system would be similar to the WRS developed during this program but would include enhanced features and a convenient means for interfacing with the DMAAC digital data base computer. The means of accomplishing the interfacing could be specified by DMAAC. Enhanced features are those which would make the proposed system the most useful to cartographers for data entry. Such features can be specified by RADC in conjunction with DMAAC with the aid of the VRS just completed.