FAA-79-24.III

REPORT NO. FAA-RD-77-177,III

SINGLE-CHANNEL VOICE-RESPONSE-SYSTEM PROGRAM DOCUMENTATION
Volume III: User's Guide

Input Output Computer Services, Inc.
689 Concord Avenue
Cambridge MA 02138

DECEMBER 1977

FINAL REPORT

Prepared for
U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Systems Research And Development Service
Washington DC 20591

PREFACE

The developmental work summarized in this final report was carried out by Input/Output Computer Services, Inc., under contract to the United States Department of Transportation, Transportation Systems Center (DOT/TSC). The research was sponsored by the Federal Aviation Administration (FAA) and represents the first phase in the Flight Service Station (FSS) Automation Program to provide preflight weather briefings to the aviation community via computer generated Voice Response. The system described in this report provides for a single-channel Voice Response System (VRS) to be used in evaluating Adaptive Differential Pulse Code Modulation (ADPCM) speech compression techniques and the man/machine communications interface for a real-time pilot self-briefing system.

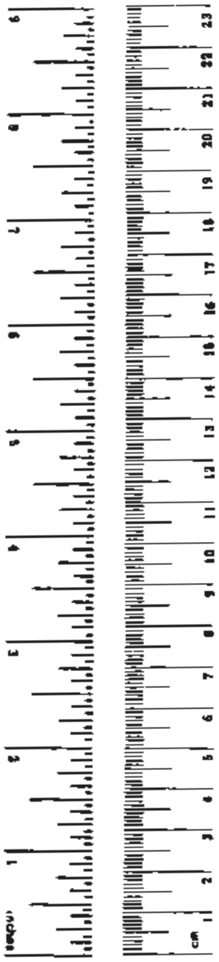
The work reported here was completed under the direction of the TSC Program Manager, Paul D. Abramson, the Technical Monitors, John Sigona and Bruce E. Ressler, and by John F. Canniff. Carey Weigel of the FAA provided overall program guidance.

This manual describes the use of the Voice Response System developed under the contract DOT/TSC 1107. An overview of system operations is given followed by detailed instruction manuals for each of the programs.

METRIC CONVERSION FACTORS

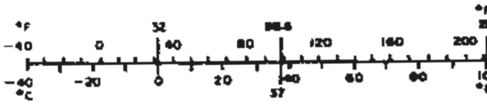
Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find
		LENGTH	
mm	millimeters	0.04	inches
cm	centimeters	0.4	inches
m	meters	3.3	feet
m	meters	1.1	yards
km	kilometers	0.6	miles
		AREA	
cm^2	square centimeters	0.16	square inches
m^2	square meters	1.2	square yards
km^2	square kilometers	0.4	square miles
ha	hectares (10,000 m^2)	2.5	acres
		MASS (weight)	
g	grams	0.035	ounces
kg	kilograms	2.2	pounds
t	tonnes (1000 kg)	1.1	short tons
		VOLUME	
ml	milliliters	0.03	fluid ounces
l	liters	2.1	pints
l	liters	1.06	quarts
l	liters	0.26	gallons
m^3	cubic meters	35	cubic feet
m^3	cubic meters	1.3	cubic yards
		TEMPERATURE (exact)	
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature



CONTENTS

Section Page

1. VRS USER's MANUAL

1.1 OVERVIEW

The Voice Response System (VRS) is a software system to provide voice output on the PDP-11 computer. The system consists of a large data base containing digitized human speech and a colleotion of programs to maintain the data base and to vocalize files of ASCII text.

The data base consists of a dictionary of text names and files of digitized speech or utterances associated with names. One file may be associated with several names, such as a proper name and its standard abbreviations. One of the names is designated the "main name" while the others are called synonyms. This distinction is required only for maintenance; lookup of words, as performed in vocalizing text files, pays no attention to this distinction.

The text names in the dictionary can contain imbedded blanks. This feature permits phrases which can be stored as one continuous utterance. The lookup algorithm used can "look ahead" on the text input to see if the next few words match a single entry or if each matches a separate word.

The dictionary is stored alphabetically to permit rapid lookup. As the system must support a dictionary of up to 4,000 entries, a simple linear search operation could not operate with sufficient speed.

1.2 VRS PROGRAMS

There are three programs in the VRS. The first is an editor to correct or update the dictionary and to add new entries. The second program is used to process human speech and separate the speech into individual utterances. These individual utterances are associated with the new entries created by the editor. The last program is linked to another computer via a serial line interface. It obtains ASCII text via this line and vocalizes this text by using the dictionary created by the above mentioned programs.

2. SYSTEM REQUIREMENTS

The VRS is designed to operate using the DEC PDP-11 series minicomputer and the RT-11 operating system. The minimum hardware configuration needed to use the VRS is:

PDP-11 processor with at least 12k of core memory

One RK03 cartridge disk drive

Console terminal such as a standard TTY

AD01 analog to digital converter

One 6 khz. external clock for the above

AA11-D digital to analog converter

Audio input and output hardware including a microphone and preamplifier (with an output of at least ±1.25 volts), an amplifier and a speaker

In addition, the last program requires a DL11 interface to a serial line. The program is linked to a Honeywell H516, but will operate with any computer obeying the line protocol (described later).

3. USING SYSTEM

All programs operate under the RT-11 monitor and can be run from the keyboard monitor. Each program contains its own command string interpreter to perform its functions. The individual programs are described in the following sections. Figure 1 illustrates the overall flow of system operations and should be referred to while following the manual.

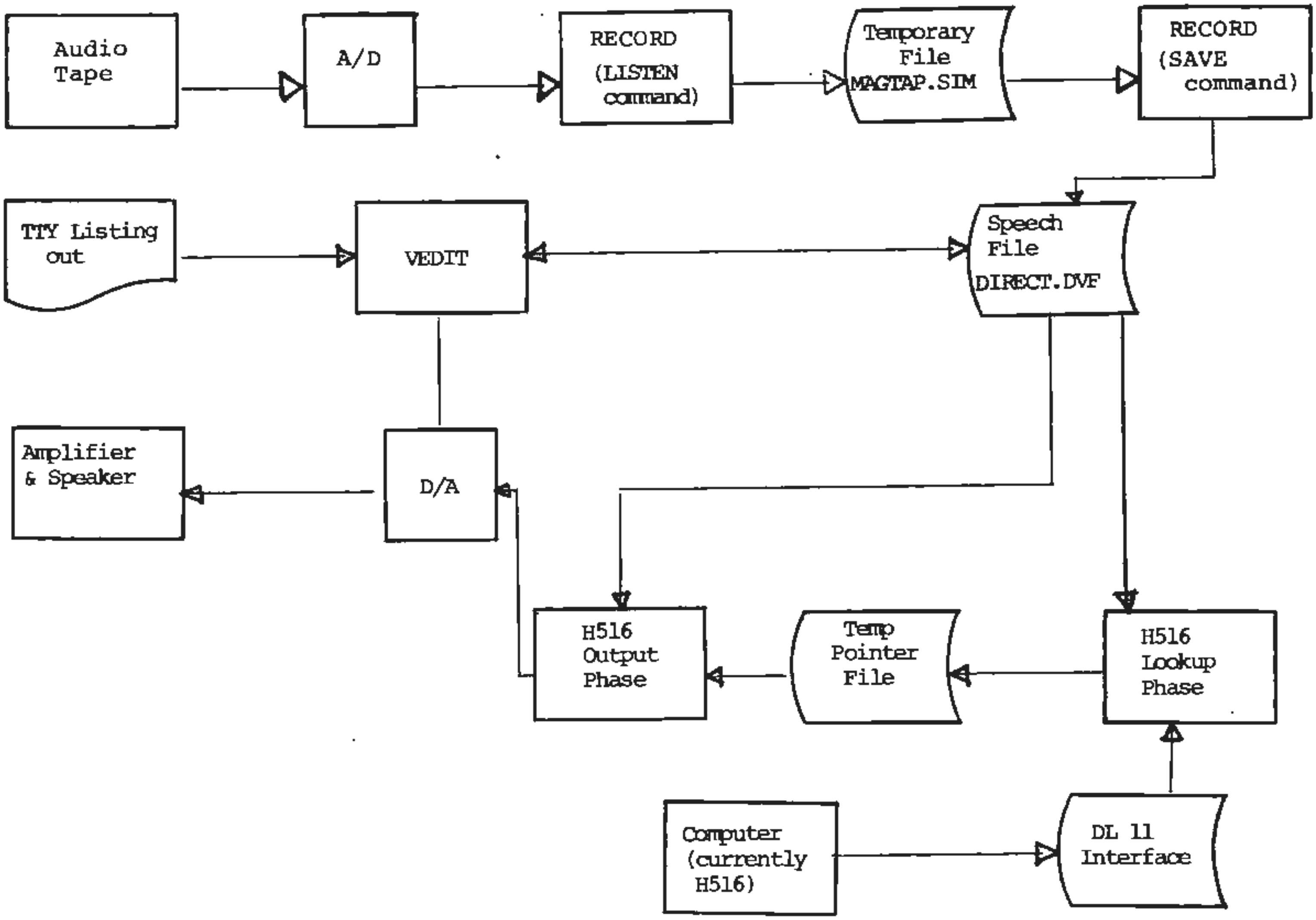


FIGURE 1. OPERATIONS FOR USING VOICE RESPONSE SYSTEM.

4. EDITOR

The VRS editor, called VEDIT, is used to maintain the dictionary for the VRS. VEDIT has most of the capabilities of a standard text editor, including commands to insert or delete dictionary entries, rename entries, or list entries and other file information.

The editor is begun ty typing "RUN VEDIT" to the RT-11 keyboard monitor. Upon entry, VEDIT attempts to open the file DIRECT.DVF which contains the dictionary and digitized speech for the VRS. If the file is not found, a new file DIRECT. DVF is created, and a message is printed indicating that file creation occurred. VEDIT then types a "?" indicating that it is ready to accept commands. When a command is entered, the program will perform the operation. A "?" will again be typed when it is finished with the operation.

4.1 COMMAND FORMAT

Commands for VEDIT are of the following format:

```
  Command/S:N   arglist
```

"Command" is the text name of the function to be performed. It is up to 6 characters long. However, only enough of the name need be typed in to provide an unambiguous match to a single command. For most commands, this is a single character.

"S" is an optional extension to the command. It is a single character used to select an option in an operation. "N" is an optional numeric value to be provided with the switch. Only two commands use this value. In most cases no value is needed.

"Arglist" is the operational list of arguments. As mentioned in the opening section, dictionary entries may contain blanks. Therefore, arguments must be separated by a non-blank separator. VEDIT uses the following separators:

```
 ,   .   ?   !   /   ;   :
```

4.2 SPECIAL CHARACTERS

A number of special characters are provided to assist command entry. These are:

CONTROL C - an RT-11 convention. If VEDIT is waiting for typed input, a CONTROL C will cause the system to return to the keyboard monitor. If an operation which does not require input (such as a listing) is occurring, a CONTROL C will not return to the monitor until the operation is completed. However, typing two successive CONTROL Cs returns the system to the keyboard monitor regardless.

> Note: When CONTROL C is used to exit, the dictionary is not properly updated on program exit. The user is cautioned, therefore, against using CONTROL C for exiting the program. This is especially true of exiting during commands which modify the dictionary. The commands "TALK" and "LIST" (described later) are the only commands that it is possible to use CONTROL C to exit without damaging the dictionary.

RUBOUT - echoes last character typed and removes it from the command string.

CARRIAGE RETURN - indicates the end of a typed command. VEDIT will then begin to process the command.

LINE FEED - treated as a carriage return.

"←" (BACKARROW) - causes character immediately following the BACKARROW to be ignored. Using this character before a carriage return permits extending a command over more than one line on the input device. BACKARROW is a shift letter "O" on a TTY model 33.

CONTROL LETTER O - causes printout to stop on the teletype until either another control letter "O" or until the next command is typed. Output which would go to the teletype is flushed.

> Note: The software producing the output continues to run as if it were outputting. Therefore, execution of a command may continue for a time after a control "O" is typed.

CONTROL U - deletes the current input line and echoes as ↑U followed by a carriage return at the terminal. (The current line is defined to be all characters back to, but not including, the most recent line feed or carriage return.)

CONTROL Z - as CONTROL U, but deletes entire command even if it was extended using a SHIFT O character. To distinguish more clearly a "%" is echoed at the beginning of a new line.

4.3 WILD-CARD OPTION

Some commands, such as LIST, do not require specific arguments, but rather can be applied to groups of dictionary entries. This is permitted by a wild card option which permits certain characters to match any character. Specifically, the following characters can be used in arguments for commands permitting the wild card option.

> " " replaces exactly one character in a match. For example "A$$" will match all 3 letter words beginning with the letter "A".
>
> "*" replaces any number of characters. For example "A*" matches all words beginning with A (including the single character "A"). "*A*" would match any word with an A in it at all.

The wild card option introduces some idiosyncrasies into the operation of some commands. A command, which permits wild card and multiple arguments, is usually carried out entirely for each argument, a pass being made for each argument in the list over the entire dictionary. Since the dictionary is sorted alphabetically, a pass is quite rapid for each argument. Even if the wild card option is used, if the "*" or the "$" is not the first character of any argument, the number of entries which must be checked is minimal. For instance VEDIT must only check words which begin with "AL" to match the argument "AL*Y".

However, if any argument begins with an "*" or a "$", it is not possible to cut down the number of checks;

a simple linear search must be performed. In such a case the program makes only one pass over the entire dictionary comparing each entry with all the arguments at once.

The effects of this on the user are twofold. First, any arguments which need not begin with an "*" or a "$" should not. In a dictionary of 4,000 words, the argument "*AL*" can cause the associated command to execute several hundred times slower than the argument "AL*". Second, if a single dictionary entry matches more than one argument, the command will be applied to the entry each time a match occurs. However, if any one argument begins with an "*" or a "$", only one pass is made over the entire dictionary and so the command is applied only once. For example "LIST AL*, AL*Y" will list the word "ALREADY" twice. "LIST AL*, *AL*Y" will list "ALREADY" only once.

4.4 COMMAND SUMMARY

The following commands are available in the current version (Version 02-04) of VEDIT:

KILL - reinitialize entire dictionary to empty.

BYE - close file DIRECT.DVF and return to RT-11 keyboard monitor.

INSERT - create new entry in the dictionary with a main name and a number of optional synonyms.

ENTER - as INSERT but is repeated for each new line until an empty line is entered.

SYNON - create one or more synonyms for a main name.

RENAME - give a new text name to an entry in the dictionary.

DELETE - remove a dictionary entry.

LIST - obtain a printout of the dictionary contents.

TALK - vocalize the dictionary contents.

GARBG - compress all free storage into a single contiguous area.

4.5 DETAILED COMMAND DESCRIPTIONS

4.5.1 K(ILL) - COMPLETELY INITIALIZE FILE SYSTEM.

Causes dictionary to be set to empty. When command is typed, the message "ARE YOU SURE?" is printed. Replying "YES" to this causes the dictionary to be initialized. The message "ENTIRE DICTIONARY INITIALIZED" is printed. If any other reply is entered the command does nothing and the message "COMMAND ABORTED" is printed.

Switch options - none.

4.5.2 B(YE) - CAUSES THE PROGRAM TO RETURN TO THE KEYBOARD MONITOR.

Switch options - none.

4.5.3 I(NSERT) - ARG 1, ARGLIST

Used to create an entry in the dictionary. The entry is created with an initial size of zero. The first argument, arg 1, is used as the main name (described in Section 1.1) while the remaining optional arguments in the arglist are used as synonyms. None of the names may already exist at the time of execution of the command. If an error occurs because a particular argument name already exists, all arguments to the right of that argument are ignored, and the argument itself is printed. The arguments to the left are inserted as above. Wild card option is not permitted.

Switch options -

Default - create file with a size of zero as described.

/S:n - create a file of silence. "n" is used to provide the size of the file. It should be an octal number between 1 and 10. This number is multiplied by roughly 90 milliseconds to produce the silence interval.

4.5.4 E(NTER) - TAKES NO ARGUMENTS.

ENTER causes VEDIT to enter what could be called "insert mode". Each of the lines following the command is treated exactly as the arguments to an INSERT command. A new dictionary entry is created for each line typed. The command is terminated by either an error, or by typing a line feed as the first character in a line.

Switch options - exactly as INSERT.

4.5.5 S(YNON) - ARG 1, ARGLIST

Used to create one or more synonyms to a dictionary entry. Arg 1 is an existing dictionary entry, but it need not be a main name. Arglist contains one or more non-existent names which become synonyms to the first name. As in INSERT, if an argument causes an error, only arguments to the left are processed. Those to the right of the argument are ignored.

Switch options - none.

4.5.6 R(ENAME) - ARG 1, ARG 2

Used to rename the dictionary entry from arg 1 to arg 2. Arg 1 is any name existing in the dictionary. Arg 2 must not exist. As a result of rename, arg 1 is removed from the dictionary and arg 2 is created. All the properties of arg 1, such as associated synonyms of main name, are transferred to arg 2. If either argument is illegal, the entire command is ignored.

Switch options - none.

4.5.7 D(ELETE) - ARGLIST

Delete all dictionary entries whose names match any name in the arglist. The wild card option is permitted in this command. If no deletions take place, such as when no entries in the dictionary match any arguments, the message "NO FILES DELETED" is printed.

Switch options:

/M - main. Delete any dictionary entry which matches any argument. For any match, whether it's a synonym or a main name, the entire entry, including all synonyms, is deleted. If the file is non-empty, its used storage is returned to free storage.

/C - confirm. Does not delete a file without first getting confirmation from the user. When an entry matches an argument in the arglist, the entry will be printed on the console terminal followed by a "?". Typing a "Y" followed by a carriage return or line feed will cause the file to be deleted as in /M. Any other reply by the user will prevent deletion.

/E - empty only. Performs the confirm operation as /C above, but only operates on files with a size of zero.

/N - non empty. As above but only for files with a non zero size.

/O - omit main names. Performs as above but only operates on synonyms. Main names are uneffected.

Default - synonyms only. Exactly as above command but confirm operation is not performed.

In either of the last two options, the message "FILE NAME IS A MAIN NAME" is printed for every file name which matches the argument list, but which is not deleted because it is a main name.

4.5.8 L(IST) - ARGLIST

Print out any file names which match the arglist and also give file information. Wild card option is permitted. There are many possible information fields to list. The switch options select any or all of these fields.

The format of a full list appears as follows:

```
File size

main           blocks          last          address
    synonym
    synonym
        .
        .
        .
    synonym
free
```

The "file size" field is a one line message telling the number of entries in the dictionary. The "main" and "synonym" fields are usually printed for any match. That is, if either the main name or the synonyms match any argument in the arglist the entire block of names and synonyms is listed. The "blocks" field give the number of 256 word blocks of storage used by the file, while "last" is the highest word written in the last block of the file. The "address" is the offset of this file in blocks from the beginning of the field DIRECT.DVF. If "blocks" is zero, then "address" and "last" are meaningless.

Finally, "free" is a two line message pertaining to free storage remaining in DIRECT.DVF. Note that all numeric information is listed in octal.

Switch options:

Default - list only the "main", "synonym" and "blocks" fields for the files matching. If both the main name and/or any of the synonyms of the main name match the arglist, the entire set of information is repeated for each match.

/A - "ALL". List every possible field. While a match occurs if either the main name or synonyms match the arglist, the "/A" switch prevents the matching entry from being printed more than once per argument. However, if a match occurs with more than one argument in the arglist, the printout is repeated for each matching arg. (See Section 4.3.)

/B - list only the main and synonym fields. As in the "/A" option, repeated listing of files is prevented.

/E - same as "/B" but will only list files with a size of Ø. Note that the command "LIST/E*" will list all files to be read to the RECORD program (Section 5).

/H - list header field and information as in "default".

/N - same as default, but non-empty files only. Also prevents repetition of file listing as outlined in /A option.

/S - "simple" lists only "main" field but does it for any field name which matches, whether it is a main name or a synonym. Wild card option is allowed for all switch options.

4.5.9 T(ALK) - ARGLIST

TALK provides audio output of any files which match the arglist. Empty files are ignored, however, silence files will perform as other files. Therefore a silence interval can be checked for proper duration using TALK.

Switch options:

Default - "speak" each entry matching an argument once.

/R:n - repeat. "n" provides a number of times each matching entry should be repeated. The repetition is on an argument by argument basis. That is for a repeat of 10, the first entry matching the arglist is repeated 10 times, then the second and so on. Note that since there is a fixed "dead time" for command setup in the execution of "TALK", R with a large value of "n" is recommended for timing pause intervals, to divide this dead time over each repetition.

4.5.10 G(ARBG)

"Garbage Collect". This command is used to compress all free storage entries into a single, contiguous free storage entry. This command may become necessary if a large number of "DELETE's" are performed.

Important: As much of the contents of the file DIRECT.DVF must be read into core, and recopied onto the file, the user should be confident of the system hardware integrity before performing this command. An error while performing this command is generally fatal, destroying the file DIRECT.DVF.

Note also that when functioning properly the system may require several minutes to perform the operation, during which time the processor will appear idle. The user is warned against the temptation to exit this operation with a control C as again the file DIRECT.DVF is likely to be destroyed.

The above warnings can be treated more lightly only if the user first obtains backup copies of DIRECT.DVF before using the "GARBG" command.

Switch options: None.

4.6 VEDIT MESSAGES

VEDIT provides the user with descriptive messages to aid the user in using the programs. Any error messages not described in this section will be of the format:

```
  ? error message ?
```

These messages are produced by the RT-11 operating system for conditions which caused an error exit from the program. These messages are described in the RT-11 System Reference Manual Section 2.8.

4.6.1 Error Messages

"filname" DOES NOT EXIST - "filname" appears in location in a command string where name of an existing dictionary entry should be. Check spelling and reenter command.

HANDLER FETCH ERROR - device handler DK could not be brought into core. RT-11 related system error. (Error not very likely.)

FILE CREATION ERROR - unable to create file DIRECT.DVF. Either RT-11 system error (not very likely) or insufficient (under 2,000 octal) contiguous blocks of free storage on device DK.

TYPE IN BUFFER OVERFLOW - attempt to make a command over 255 characters long.

NEEDS ARGUMENT - no argument provided for command which expects argument, or for an insufficient number of arguments for a command which requires more than one. Commands permitting wild card option do not print this message.

DISK READ ERROR - RT-11 system was unable to perform a disk read. User should consult "RT-11 Programmers Handbook" Section 9.4.30 to determine why read could fail.

DISK WRITE ERROR - as above but on disk write consult Section 9.4.47. Either read or write likely to be a hardware rather than software error.

GARBAGE COLLECTION DISK ERROR - Printed if either of the above errors occur while performing the "GARBG" command.

BAD SWITCH VALUE "val" - Printed if numeric switch value is incorrectly entered. If the switch value entered contains a character other than 0-7 the illegal character and remainder of the numeric argument is printed as "val". If the entered number is a valid number but outside the range for the particular command, the first argument following the switch value is printed.

ARE YOU SURE? - See KILL command.

COMMAND ABORTED - See KILL command.

ENTIRE DICTIONARY INITIALIZED - See KILL command.

Filnam IS A MAIN NAME - See DELETE command.

NO FILES DELETED - See DELETE command.

AMBIGUOUS COMMAND: "com" - "com" matches more than one command. Retype command including more characters.

"filnam" ALREADY EXISTS - "filnam" appears in argument position where a new, non-existent file name should be. Check spelling and re-enter.

ILLEGAL SWITCH - syntax incorrect for use of switch option. Re-enter command.

BAD ARGUMENT - Argument to command is invalid or argument not provided to a command which expects one.

DICTIONARY FULL - Area in Dictionary reserved for text is full, or is less than needed to complete the attempted command.

NO SUCH COMMAND: "com" - "com" does not match the name of any legal command.

NO FREE STORAGE TABLE ENTRIES - (not likely error). Space reserved for free storage tables is exhausted.

INITIALIZATION ERROR - One of the operations involved in system start-up failed. RT-11 related system error, or write lock on disk.

FATAL CONSISTENCY ERROR - System related error which should not occur. In the event of such an error, save the results of the current terminal session and reboot. IOCS programming services should be informed of the occurrence and the circumstance.

4.6.2 Informative Messages

NEW FILE "DIRECT.DVF" CREATED - Editor is started and no file DIRECT.DVF exists. If attempt to create new file is successful, message printed.

VEDIT VERSION n - Inform user which version of VEDIT he is using. This message is not printed for version Ø. This document is written for version 02-04.

5. RECORD

RECORD is a program to input speech to the files in DIRECT.DVF. It is used in conjunction with VEDIT to create a Dictionary for the Voice Response System.

There are three steps to inputting the speech. First, the words to be entered are read into a microphone by an announcer. The words are obtained by using the "LIST/E" command to VEDIT (Section 4.5.8). Next, the speech processing is initialized by setting the energy threshold for the utterance editing algorithm. Finally, the utterances are processed, separated, and the individual utterances are stored in the dictionary.

As VEDIT, the program operates under the RT-11 keyboard monitor. It is started with the command "RUN RECORD". When running RECORD, the dictionary must exist in the file DIRECT.DVF. If it does not, the program will print a descriptive message and return. The program also requires 24 decimal blocks for contiguous free disk storage on device DK for every second of speech recorded.

5.1 COMMANDS

The command structure for RECORD is almost identical to that of VEDIT. The only difference is that RECORD commands do not have switch options.

5.1.1 G(AIN) arg

GAIN sets the gain on the A/D input to RECORD. The value of arg selects the voltage value which will produce the full scale digital output for the D/A. The voltage range selected should be roughly two times the AC RMS voltage level in the signal being input to the A/D to take peaks into account.

The effect of the settings are

Gain Value	Peak To Peak Voltage
Ø	±10 volts
1	±5 volts
2	±2.5 volts
3	±1.25 volts

The program is started with a gain of 2. Generally, this is correct for a standard preamp output from a tape deck when the recording level meter is reading Ø db (decibels).

5.1.2 L(ISTEN)

LISTEN is the command used to input the speech to a temporary file for later processing. This step is required because the processing cannot be done in real-time.

The steps in the "LISTEN" process occur as follows: When LISTEN is typed, the console bell is rung. At this time the words on the listing are read into the microphone or are fed into the A/D from a previously made audio tape. This continues until a second bell is rung. The speech is now saved in a file named MAGTAP.SIM.

Several rules should be observed by the announcer when speech is input.

a) Only the main names (those words not indented on the listing) should be read. These should be read in the order they appear on the listing.

b) The processing of the speech is somewhat noise sensitive. Therefore, it is recommended that the speech be input from a previously made audio tape rather than by a microphone which will pick up the computer room noise.

c) A pause of at least 1/6 second between utterances is required. While this seems an easy constraint to meet, the announcer should pay particular attention to separation between words that begin or end with phonemes such as "th" or "s" which tend to cause words to run together.

d) Often, due to limited disk capacity, the required number of utterances cannot be input at once. When the number of words input is less than the number of entries listed by the "LIST/E*" command, it is often good to run the program VEDIT to verify that the last word input was indeed saved, and that the following word still has a size of zero.

Eventually the file MAGTAP.SIM will be replaced by an actual Mag tape unit. The only effect this will have on use of the LISTEN command is to require proper set up of the unit (scratch tape, set at load point). Rule 4 will still be a good idea to follow but the capacity limit will not be as restrictive.

5.1.3 T(HRESH) - ARG

The speech must be processed after it has been recorded. The processing involves two operations. First, the recorded speech in the file MAGTAP.SIM is reduced by employing a speech processing technique known as Adaptive Differential Pulse Code Modulation (ADPCM). The details of the ADPCM algorithm are unimportant for the purposes of this manual. It is just important to note that by taking advantage of the correlation between successive samples of speech, it is possible to reduce the storage requirements to only four bits per sample.

After this compression of the speech, the utterances are separated. The separation is accomplished by computing the "energy" of the above four-bit samples. When the energy is greater than a pre-selected value for a given period of time, it is determined an utterance has begun. After the energy drops below this value (for a predetermined period) an utterance has ended.

This preselect value is input using the THRESH command. The value input must be an octal number.

There are no cut and dry rules for setting the threshold value. The behavior of the ADPCM "energy" function, as described in reference 1, volume 1 of this report, has no one to one correspondence between voltage level in the input and the value of the energy function. The value of threshold therefore must be found experimentally.

Some guidelines for determining the best value are presented here. These are only guidelines and the user may find other values better suited.

Experimentation has determined a good mean value of roughly 6,000 - 7,000. At this value words ending in stop consonants such as "t" or "p" may end quite abruptly. Some words such as those ending in "s", "th", or other unvoiced sounds may be a little prolonged. In general it is good practice to do an initial run of the program with the threshold at roughly 7,000 and then delete and re-enter utterances which are too short or too long. Those utterances which are too long can be shortened by using a higher threshold, the short ones can be lengthened by using a lower value. Some practical limits are 3,000 for a low end,as a normal signal level for background noise is rarely below this value, and roughly 12,000 - 14,000 for a high value. In the latter vowel sounds begin to be edited out. The user should perform some experimentation to obtain a feel for the effects of threshold values.

One final note concerns signal to noise ratio. Again by the same reasoning as above, no real hard rules exist. The only guideline is that if a threshold value of much above 4,000 - 5,000 is needed to even separate utterances from the noise, it is likely that the audio portion of the installation (that is microphone, tape, and filter) is too noisy to obtain a good quality vocabulary.

5.1.4 S(AVE)

SAVE actually starts the speech processing. Each time an utterance is detected or separated using the algorithm described in 5.1.3, it is stored in a permanent file and is associated with the alphabetically first empty entry in the dictionary. This continues until an end of file in MAGTAP.SIM, until the empty entries are exhausted, or until an error occurs. The program types a descriptive message indicating the cause of command completion. Possible error conditions include RT-11 disk read or write errors, and exhaustion of free storage space in DIRECT.DVF.

5.1.5 B(YE)

Causes the program RECORD to return to the keyboard monitor.

5.2 ERROR MESSAGES

DISK RECORDING ERROR - RT-11 related system error occurred in recording A/D to disk. RT-11 hardware or software failure.

ERROR IN FILE CLOSING - Disk read or write error (RT-11 hardware or software error) occurred when backing up the directory at the completion of the "SAVE" command.

DISK ERROR - RT-11 hardware or software disk I/O error, or if error occurs in "SAVE" command could also indicate insufficient storage to save the utterance being processed.

FILE SPACE EXHAUSTED - DIRECT.DVF full (no free storage).

DICTIONARY EMPTY - Dictionary has no entries under which to save utterances as they are processed.

NEEDS ARG - Argument absent from a command.

TYP BUF FUL - Attempt to make a command over 255 characters long.

NO SUCH COMMAND: "com" - "com" is not the name of any legal command.

FATAL CONSISTANCE ERROR - See Section 4.6.

END OF FILE ENCOUNTERED - End of file MAGTAP.SIM encountered while processing an utterance.

NO EMPTY DICTIONARY ENTRIES - No empty dictionary entries are available to be associated with utterances. Printed when all empty entries are used by a SAVE command.

6. H516

H516 is a program to accept ASCII input from the H516 computer and provide audio output in English corresponding to the ASCII text. It performs the following functions.

1 - Maps single utterances to single ASCII strings. These utterances may be single words or entire sentences.

2 - Through the synonym feature, described earlier, it can map several different ASCII strings to the same utterance.

3 - Speaks numeric data in natural English. For example 395 is spoken "three hundred ninety five".

4 - Several utterances may be spoken for each occurance of a single string; for example "HK" can be spoken as three separate utterances "HAZE AND SMOKE", when the proper entries are made in the dictionary.

5 - Punctuation is converted to pauses to provide a more natural sounding output.

6.1 LINE PROTOCOL

The program communicates with the H516 computer via a single 9600 baud communications line. The protocol used is line at a time with no provision for error checking. In the protocol, the PDP-11/20 is the "master", that is, all data from the H516 is sent only in response to a request from the 11/20.

The format for the request is a single ASCII character from the 11/20. The H516 responds with a line of text ending in a carriage return (ASCII 15 octal). The first character is an ASCII numeral indicating channel number. In this single channel system, that character is ignored.

6.2 DATA

The program uses two data bases: the ASCII data sent from the H516, and the vocabulary created using the programs VEDIT and RECORD. For proper operation of the program, several constraints must be followed concerning the data.

6.2.1 Vocabulary Constraints

Entries in the vocabulary can consist of any ASCII characters except the following separator characters:

; : . , / ? ! - carriage return - line feed

Any other characters in the ASCII set are permitted. In practice, however, it is recommended to restrict entries to the following subset of ASCII:

A-Z, 0-9, Space, #, @

In addition, "@" has special significance (explained later), therefore should not be used except in the special case.

The only entries which are actually required to permit the program to run are the numerals 0-19, 20, 30,...90. The way numbers are parsed causes program failure if these entries are not in the vocabulary. Other entries which are important follow. These entries are required for proper performance of the program. However, the program will still operate without them. The entries are:

The Alphabet - A-Z
The words "HUNDRED" and "THOUSAND" used in parsing and speaking numbers.
Any characters which produce an audible response when encountered alone.

6.2.2 Contractions

The vocabulary must contain special entries to expand contractions into several utterances. This constant applies only to the vocabulary, i.e., no special characters identifying a string as a contraction need be placed in the data from the H516.

Contractions are expanded by providing synonyms to the words into which the contraction expands. The synonyms are of the following format.

```
cntrct@n
```

"cntrct" is the contraction as it appears in the text from the H516.

"n" is a sequence indicator. That is, it is an extension which gives the particular synonym a particular alphabetic position with respect to other synonyms using the same contraction. By convention the digits 0 to 9 are used.

To make this notion more clear, consider the contraction "HK" which is spoken "HAZE AND SMOKE". In the dictionary "HK@0" would be made a synonym to "HAZE", "HK@1" would be made a synonym to "AND" and "HK@2" would be a synonym to "SMOKE". When the progran encounters an "HK" in the input string, it will output all utterances beginning with "HK@" in alphabetical order.

If entries such as above are in the dictionary simultaneous to "HK" with no "@" ending, the result of the lookup is undefined, but will not be a fatal error.

6.2.3 Constraints on H516 Data

The data from the H516 should correspond to the vocabulary. That is, strings from the H516 should not contain any characters which do not appear in the vocabulary.

The "@" must not appear in the H516 data. Punctuation is converted to silences of the following length:

","	-	approx.	170	millisecond	pause
"/"	-	"	170	"	"
":"	-	"	256	"	"
";"	-	"	256	"	"
"."	-	"	340	"	"
"?"	-	"	340	"	"
"!"	-	"	340	"	"

Currently, characters which are not in the vocabulary are ignored. This may change after review of system performance.

Numerals are expanded and spoken in only one way, as described in Section 6. While the exact way they are spoken may change after performance review, nevertheless they will still only be spoken in one way. If the user requires different expansion, the full expansion as it will be spoken must appear in the ASCII input.

180 Copies