

COVOX inc.

675-D Conger Street • Eugene, Oregon 97402 • U.S.A.
Area Code (503) 342-1271 • Telex 706017 (Av Alarm UD)

VOICE MASTER OWNER'S MANUAL

VERSIONS 2.19 AND 3.0

FOR COMMODORE 64 & 128 IN 64 MODE

October, 1985

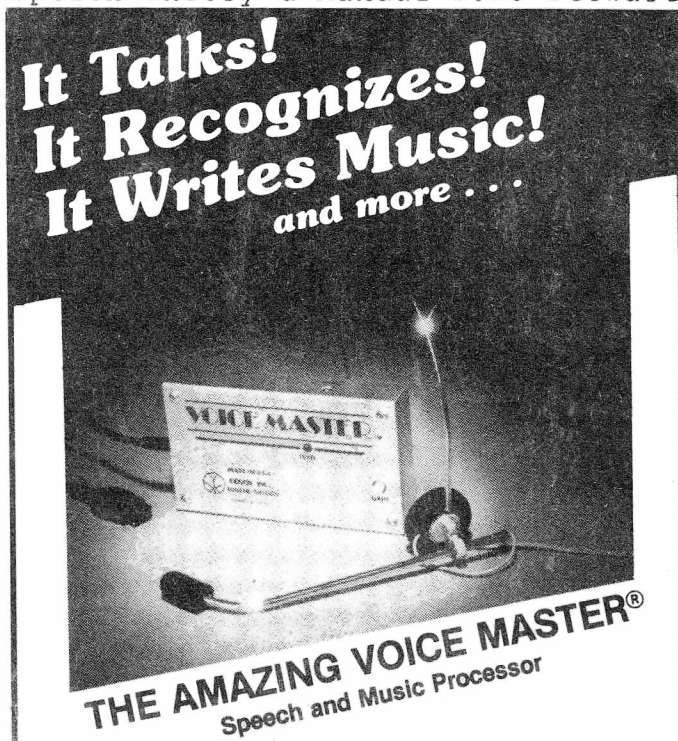
PART 1: SPEECH RECORDING AND PLAYBACK

PART 2: SPEECH RECOGNITION

Copyright 1985 by
COVOX, INC.

CREDITS

Project Leader: Bradley C. Stewart
Senior Programmer: Neddie D. Higgins
Speech Theory & Manual J.L. Stewart



BLANK

TABLE OF CONTENTS

INTRODUCTION.....	1
CALIBRATION CHECK.....	3
EARPHONE.....	5
DEMONSTRATION PROGRAMS.....	6
IMPORTANT INSTRUCTIONS FOR LIST AND RUN.....	10
.....	
PART 1: SPEECH RECORDING AND PLAYBACK.....	11
ELEMENTARY PROCEDURES.....	11
SOME CLARIFYING COMMENTS.....	15
SAVING AND LOADING VOCABULARIES.....	17
SPEED, SAMPLING RATE, LEVEL, AND SCREEN.....	19
THE CLEAR COMMAND.....	20
THE IMPORTANT TASK OF MAKING A BACKUP.....	21
LOADING VOICE MASTER FROM BASIC.....	22
THE PLAY/FIND PROGRAM.....	23
OTHER PROGRAMS AND COMMENTS.....	25
SPECIAL DATA CHECKING PROGRAM.....	26
.....	
PART 2: SPEECH RECOGNITION.....	29
PRELIMINARY COMMENTS AND SUGGESTIONS.....	29
SPECIAL KEYS AND SYMBOLS.....	30
RECOGNITION PROGRAMS.....	31
PRINCIPAL STATEMENTS.....	32
PARTITIONING.....	34
SAVING AND LOADING TEMPLATES.....	35
THE "VOICE/RECOG" PROGRAM.....	36
DEMONSTRATION RECOGNITION/PLAYBACK PROGRAM.....	38
.....	
COMMAND SUMMARY (Recording/Playback).....	40
IMPORTANT MEMORY LOCATIONS (Recording/Playback)...	42
COMMAND SUMMARY (Recognition).....	44
IMPORTANT MEMORY LOCATIONS (Recognition).....	44
SOME GENERAL COMMENTS ON MEMORY USE.....	44
.....	
APPENDIX: CASSETTE OPERATING SYSTEM FROM DISK.....	47

LIMITED WARRANTY STATEMENT

COVOX, Inc. guarantees the VOICE MASTER to be free from defective materials and workmanship for a period of one year from the date of purchase. COVOX, Inc. will replace defective parts and make repairs under this warranty when the defect occurs under normal use, provided the unit is returned to the factory via prepaid transportation. The warranty provides that examination of the returned product must disclose a manufacturing defect to be judged by COVOX, Inc. The warranty does not extend to any product which has been subject to misuse, neglect, accident, improper installation, or where the panel legends or other markings have been removed or defaced, and is given in lieu of any other warranty implied or expressed, and will not cover any consequential damages.

Information in this manual and associated software are provided on an "as is" basis. No warranty, either expressed or implied, is made by COVOX, Inc. pertaining to suitability for any specific application or commercial use. It is the purchasers responsibility to make appropriate evaluations for such purposes. COVOX, Inc. disclaims liability for direct, indirect, or incidental damages arising from the use of this product, including but not being limited to interruption of service, loss business or potential profits, legal actions, or other consequential damages.

PATENTS AND COPYRIGHTS

The software supplied with VOICE MASTER is copyrighted. It may not be copied, reproduced, translated, or reduced to any readable medium or code for other than personal use without prior written permission of COVOX, Inc.

The hardware/software system comprising the COVOX VOICE MASTER is subject to existing patent applications. Unauthorized duplication for commercial purposes or to otherwise avoid payment of appropriate royalties or license fees will be deemed to be a violation of proprietary rights under patent and trademark laws.

The names COVOX, VOICE MASTER, and VOICE HARP, and the COVOX "logo" are registered trademarks and are the property of COVOX, Inc.

INTRODUCTION

This manual is in three parts. Part 1 concerns speech recording and playback. Part 2 covers word recognition which operates in conjunction with software for speech digitization as in Part 1. Part 3 is concerned with voice production of music with software separate from that for speech. Music is discussed in a separate manual.

Software for VOICE MASTER is supplied on disk. A Tape cassette version is not available from COVOX, Inc. However, VOICE MASTER software does provide means for preparing a tape based system. A disk drive is required.

Pre-recorded vocabularies produced according to Part 1 of the manual may be saved on disk and reproduced on a separate Commodore 64 computer without Voice Master hardware. If a speech vocabulary is produced in this way and then sold as part of a commercial software product, we request that the COVOX name and address be given in sufficient detail to permit purchase of hardware by the software customer. With this acknowledgement, use of COVOX software for speech recording and playback is free of royalties. The VOICE MASTER is required for speech recognition and music. Software authors are advised to contact COVOX, Inc. if recognition software is to be incorporated into programs for sale.

A number of programs are provided on disk, some essential to speech recording and playback, and some which demonstrate applications. One program in BASIC is for a talking calculator. Another a talking clock. Yet another helps direct you in the creation of word templates for speech recognition. A talking and listening calculator program is present, along with a blackjack game that takes bets from your voice input and reads back the cards. Programs written in BASIC can be listed for study (with a special procedure as will be explained.) Programs in machine language also can be listed and disassembled in the usual way.

NOTE: Your VOICE MASTER has a small hole labeled CALIBRATE. A special calibration routine should be undertaken before operation; or at least the calibration should be checked. This will provide useful experience pertaining to the VOICE MASTER, including a valuable speech display program. It also will instruct on part of the installation procedure.

BLANK

CALIBRATION CHECK

Although the VOICE MASTER has been calibrated and checked prior to packaging and shipping, variations in individual computers can sometimes be large enough to warrant a check of the calibration. Proceed as follows:

1. Connect together the parts of your computing system as described in the Commodore owner's manual.

2. Turn on the computer without the disk in place.

3. Plug the VOICE MASTER into the rear control port (port number 2).

4. Plug the microphone (^{RED}~~black~~ phone plug) into the jack called MIKE. (You can plug in the earphone next to this, at EAR, but this will not be used for the present.)

5. Speak into the microphone in a normal voice and adjust the GAIN control (starting from minimum, or full counter clockwise) so that the LEVEL light on VOICE MASTER flashes on speech peaks. Use a sustained vowel sound such as "ah" for this adjustment. We suggest placing the microphone at and just above the lips so that nasal sounds are able to get to the microphone.

6. Now install the VOICE MASTER disk. The computer is presumably in the READY mode showing a flashing square cursor.

7. Issue the keyboard command

LOAD"BAR",8,1 (and press RETURN key)

The name of the program you are LOAD(ing) is "BAR". The number 8 signifies that you are loading it from the disk drive which (usually) has reference number 8. The number 1 tells the computer to put the program in the same place in memory that it came from when originally produced by the programmer.

You are not yet done. After the READY prompt reappears on the video (or TV) display, then type in another command as

SYS 20480 (and press RETURN key)

The SYS command tells the computer where in its memory the program is supposed to begin. Machine language programs are generally started with SYS to some location.

Now you see before you a display of dancing vertical bars which indicates the range of sounds that you are making into the microphone. A number also appears in the lower right corner of the display. The information contained in these bars is used in part for word recognition. Now pay attention to the vertical bar to the right of the display. The height of this bar, and the value of the number in the lower right corner, represent amplitude, or volume. When you are not speaking, this bar should not appear at all, not even a small amount. And the number should have the value 0. When you produce a soft speech sound, the bar should appear. If the bar appears when there is no sound, or if it takes too much voice power to make it appear, then you must undertake the following calibration:

Use the small "tool" provided with VOICE MASTER (or a small blade screwdriver) to turn the control in the CALIBRATE hole so that the bar just disappears when you do not speak, and is "reasonably" out of sight in the presence of background noise. Or, you may find it easier to refer to the number for this adjustment, setting the number to zero midway between of adjacent numbers. (On one side, the display may wrap around so that the bar becomes very large. Interpret this as meaning a negative reference number.)

Now go back and recheck the GAIN. When you say "ah", the amplitude bar should flicker about the level where the color of the bar changes from green to red on a color monitor, or shading changes on a monochrome monitor. When you produce a nasal such as "m", the bar should flicker near the base line, about the numbers 1 to 4 or 5. If not, then readjust the microphone.

While you have the "BAR" display, observe behavior of the dancing pattern. Try sounds like "eee", "ah", and "ooo". The graph plots sound frequency periods with the shortest periods (highest frequencies) appearing on the left. It is a little like a real time frequency spectrum analyzer, but with the frequency scale opposite from the conventional direction. The "BAR" display is normalized so that a sweeping pure tone of constant amplitude will cause all bar segments to achieve (approximately) the same maximum value.

Also look at the next to last bar on the right. Its color is light blue on a color monitor. The height of this bar measures your fundamental voice pitch. This part of the display might be useful in speech training.

Proper calibration is very important! If done wrong, recorded speech will contain background noise, or be excessively distorted. In speech word recognition, this noise or distortion can severely reduce recognition accuracy and prevent the computer from determining where a word starts and where it stops.

In order to get out of "BAR" and get back to the ready blinking cursor, press the space bar. No other key will work (except pressing both RUN/STOP and RESTORE keys at the same time).

Note: We will henceforth leave off the "(and press RETURN key)" after a command. This requirement is "implied". Pressing RETURN signals the computer to accept what was typed. No RETURN means that no command (or statement) has been entered, and the computer will simply wait for you to do something. (There are some exceptions to this rule, one being when the computer waits for speech input as will later be discussed.)

The computer does not understand a command if it is spelled wrong. If you type a keyboard command incorrectly and RETURN, you may get an error message. Simply retype. Or, before RETURN, backspace one or more times with the INST/DEL key on the upper right of the keyboard and type again. (The computer may accept an incorrect command which it is typed in as a numbered statement. But the error will be detected when the program is listed or run.)

If for some reason the computer "locks up" so that nothing you do with the keyboard seems to get you out of a predicament, don't be bashful. If RUN/STOP with RESTORE fails, then take the ultimate step (with the C64) and simply turn everything off (first removing the disk). Then turn on the computer and install the disk. Turning off the computer destroys all of the data and programs in memory that you are currently working with. The prudent programmer periodically saves work, finished or not, on the disk. The C128 is equipped with a reset button that can be pressed so that everything need not be turned off. Some C64's have been "retrofitted" with a reset button as well. With reset, power need not be removed, but BASIC programs and data may still be lost (although some machine language programs may remain). The C128 in the 128 mode, but not in the 64 mode, can be reset in a way which preserves programs.

EARPHONE

If your headset microphone is also equipped with an

earphone, you can have silent listening to voice or other sound response from the computer. The earphone jack on VOICE MASTER is next to the microphone jack. But no audio is available at this jack unless a cable is connected from the phone jack on the opposite side of VOICE MASTER, labeled EAR IN, to the computer. Two audio cables are provided for this purpose. One is for use with a TV monitor and the other works with a video monitor. The TV cable has a "DIN" connector (large round) at one end which goes to the video output port on the rear of the Commodore 64. A video monitor is usually supplied with a cable that connects to a "DIN" socket on the back of the computer, and at the other end has separated leads with RCA type male connectors (usually 2 or 3). The RCA female connector at the end of the cable supplied with Voice Master goes to the male RCA connector that would otherwise plug into the audio input to the video monitor. If the monitor does not have audio capability, then the earphone will provide the needed sound output.

DEMONSTRATION PROGRAMS

A BASIC program called "MENU" allows selection of several programs which illustrate the various capabilities of VOICE MASTER. Load as

LOAD"MENU",8

and then the program must be RUN. Eight programs are listed and a ninth selection gets you back to the computer READY state. All programs are menu driven with instructions for use provided in the pertinent sub-menu displays. "MENU" loads in the all-important machine language program which handles speech recording, playback, and word recognition. There are other programs on the VOICE MASTER disk which do not handle everything at once (i.e., recognition only, or speech playback only), and these will be shorter in length and quicker to load. But for the several demonstrations given here, there is no particular advantage to saving memory. If you should wish to load in the all-inclusive program without the demonstration program overlay, then LOAD"VM2.1R",8,1. Then SYS49152. Then NEW.

A quick introduction to VOICE MASTER is given with "DEMO" which is loaded with selection number 1 on the main "MENU". (The program "DEMO" can be loaded directly, but it will not LIST or RUN without "VM2.1R" in place.) Selections from the sub-menu utilize the function keys f1, f2, etc. (with capital F for the C128). Upon selection of "DEMO", you will not only get a menu listing, but also a voice

saying "COVOX VOICE MASTER..TER..TER..TER" where the "TER" part is rather like an echo. You may call this "institutional advertizing". In fact, all of the demonstration programs may be viewed in this light, although they are a necessary part of teaching a person how to best employ VOICE MASTER in practical programs.

Press the f1 key once and you are presented with a preparatory message. Now have your microphone ready and press f1 a second time. Then say something. But do not pause or the recording will stop. The screen may go blank after pressing f1, but recording does not (normally) start until you start talking. The screen comes back on when the recording stops, which happens when a short period of silence occurs.

Now that you have recorded something, play it back by pressing the f3 key, with the volume on the TV or video monitor turned up enough for your needs. You can experiment with this again and again. Try a short "hello" and then press f5. You will get an echo effect. Try pressing f7 and you get your "hello" back at several different speeds, like a slowed or speeded tape recording.

A kind of stuttering can be achieved during playback. After pressing f3 and before the playback ends, press the space bar or the "Q" key for reset to the start (as rapidly and repeatedly as you want). A variety of interesting rhythm effects can result.

Pressing f8 for "SPECTRUM DISPLAY" gets the same "BAR" display previously discussed.

Press f2 and you will be asked to input 4 words which create recognition "templates". You will then be able to speak one of the words into the microphone and the computer will respond by printing that word on the display.

In order to get back to the main "MENU", press the "Q" key, meaning "quit". If you want to get back to the READY state, press RUN/STOP.

The program "CLOCK" on the main "MENU" comes with selection number 2. A sub-menu provides instructions. Start by entering the time. Of itself, this provides a real time clock for the C64. Then enter a vocabulary. You are prompted to enter the correct words. Finally, you can operate the clock and get time with a voice. An alarm can also be implemented with a suitable "get up" message.

Selection 3 is "CALCULATOR". The idea is to record

numbers and mathematical operators so that you will have a four function calculator. You will be prompted to record the proper words. As you press the various number and symbol keys, the computer talks out the operations and the answer.

Selection 4 gives "RECOG-DEMO". You are asked to input words so as to make templates for several colors. The program has you go through the process twice so that the final template for each word is an average of two utterances. Template making is sequential--no keys need to be pushed for each individual input word. Then you are asked to input the names of the colors, using the fl key to start each word. When you are done, the program goes direct to the recognition mode. When you speak a color, the proper word and screen color (if you have a color monitor or TV) will be presented. You can of course make a substitution error. The computer can refuse to make an identification with a code number 254 or 255. This means that two of the templates yield too close a match to be called, or that no match is good enough. The error messages, and methods for changing error criteria, are described later in this manual.

A variation of "RECOG-DEMO" is to make templates for the colors in one language, say Spanish, with the voiced response in English. Sort of a Spanish-English dictionary, and possibly useful in language instruction. Another possibility is to produce numbers instead of colors--the color screen prompts need not be followed.

"BLACKJACK" is a real blackjack game! You are asked to make templates for the various words you employ in playing the game. There is a dealer voice that has been pre-recorded, but you can make your own dealer voice if you wish. The words you use are "hit me", "stand", "stop", the nine digits, "cards", "bet", "yes", "no", and "erase". You start the game by saying "yes" to the prompt. You are given \$20 to start with. You must first tell the computer how much you wish to bet, such as "four", or a larger sum such as "one...two" for "12". "Erase" corrects a betting amount error, if one is made. If no number appears in response to you voice, repeat. After saying the amount you wish to bet, say "bet". The computer responds by dealing you cards and speaking out what the cards are. You then say "hit me" or "stand". If you don't "bust", the "house" deals its own cards and the outcome of the game is announced. Saying "stop" returns to the main "MENU". The RUN/STOP key returns you to the READY state. At this time, you can re-TRAIN or re-LEARN certain words. (Necessary details are discussed later in this manual. The word,

along with the corresponding index number, can be seen in the DATA statements starting at line 269 in the BASIC program. You may also wish to save the speech patterns and templates on a formatted disk. (You can't save to the VOICE MASTER disk because it is write protected.) After making adjustments, you can continue the program by typing CONT and then the RETURN key. If the computer can't continue (maybe because you made a syntax error), then type CLR:RUN11 and RETURN. The game will restart but you won't have to re-enter the vocabulary. If you bet over your limit, you will have to start over. You can review the cards played by saying "cards" while placing your bet. Pressing the up arrow key (not the cursor key) puts the game on hold.

Part of the "BLACKJACK" program is "VOICE-CREATOR" which is a BASIC program that allows you to use your own voice for the dealer. You will have to use a different formatted disk in order to create a new vocabulary for this purpose. Save both "VOICE-CREATOR" and "BLACKJACK" on the formatted disk (both BASIC programs). Then save "VM2.1R" using the "BACKUP" program that comes with your VOICE MASTER. When you RUN "VOICE-CREATOR", "VM2.1R" will be loaded into memory and you will be prompted to speak a list of words. After that, you will have several editing options. You can speak any phrase and re-enter ones that don't sound right. When satisfied with your vocabulary, press the F8 key (shift F7) and your voice patterns will be saved on the disk as a file called "VOICE". Then you can run "BLACKJACK".

"CALCULATOR-2" is a variation of "CALCULATOR". But your word commands replace pressing the keys. Just say, for example, "three..plus..six..equals.." and the computer will tell you the answer. Load and run this program as with any other program written in BASIC. You will be prompted to enter the nine digits (zero through nine), "point", "plus", "minus", "times", "divided by", "equals", "clear", "delete", and "stop". When the computer prints "What shall I calculate?", then it is ready for voice command. Saying "clear" will clear the screen, "delete" will erase the previous entry, and "stop" will return to the main "MENU". The READY state is reached with the RUN/STOP key and then you can re-TRAIN, re-LEARN, or save your speech patterns and templates (but not with the write-protected VOICE MASTER disk). Type CONT and RETURN in order to continue the program. Pressing the up arrow puts the calculator on hold.

Choices 7 and 8 on the main "MENU" select programs relating to music. These are discussed in the separate

manual on the "VOICE HARP".

IMPORTANT INSTRUCTIONS FOR LIST AND RUN

When you first turn on the computer, it will understand all of the commands defined for Commodore BASIC. An incorrectly typed keyboard command will result in a syntax error when you RETURN. A command in a numbered BASIC statement will be accepted when originally typed in with RETURN. But when LIST(ed) or RUN, the computer will stop at the offending line number with a syntax error.

VOICE MASTER programs define a number of new BASIC commands which are "wedged" into Commodore BASIC. But Commodore BASIC will not understand these commands unless it is augmented with the appropriate VOICE MASTER program.

Stated simply, you cannot LIST or RUN a BASIC program that contains VOICE MASTER commands unless the proper VOICE MASTER program has been loaded into computer main memory so that it can augment Commodore BASIC. (A program can be SAVE(d) or LOAD(ed) without a VOICE MASTER PROGRAM.) The program called "VM2.1R" handles recording, playback, and recognition. "VR1.0" handles recognition only. "VR1.0" will not allow VOICE MASTER commands pertaining to recording and playback because it enhances Commodore BASIC only to the extent of commands relating to word recognition. There are two machine language programs which implement playback and recognition that do not use wedges. They may be embedded in a user program without first having to load in one of the VOICE MASTER programs with wedges.

PART 1: SPEECH RECORDING AND PLAYBACK

ELEMENTARY PROCEDURES

Now you can start your own programming. We begin with some easy direct keyboard commands. If required, details will be given later on in this manual. But for most applications, what is said in this section may be enough. We start with the assumption that you have loaded in the proper machine language program. If your interests are only in recording and playback (no recognition), then load in "VM2.1" or "VM2.1N". These two programs differ in that the first contains a special software routine which loads in a saved speech vocabulary from disk faster than the standard method that is incorporated in "VM2.1N". The fast method is not applicable if the disk drive is not model 1541 or if the power line frequency is not 60 hertz. Also, the fast method is not applicable for loading from tape cassette. An alternative is to load in "VM2.1R" which carries along software for word recognition as well as recording and playback. This composite program does not contain the fast loading routine. Load one of the three programs (VM2.1, 2.1N, 2.1R) as

```
LOAD"VM2.1N",8,1
SYS 49152
NEW
```

where SYS and NEW should not be issued until the READY prompt indication is given. (Decimal 49152 is C000 in hexadecimal.) The alternative programs load in the same way--just change the name of the program. The "NEW" in the above resets pointers in BASIC so that you will be able to enter commands and statements.

If you happen to have one of the "MENU" programs in memory, you can strip everything out of memory except for "VM2.1R" by typing "NEW" and then "CLEAR". The "NEW" removes BASIC statements from the ongoing "MENU" program so that they will not interfere with your entries. The "CLEAR" is a VOICE MASTER command that clears out whatever word vocabulary happens to be in memory. We generally suggest, however, that the desired program directly be loaded in.

Note: It is sometimes easy to overlook the need to have one of the VOICE MASTER programs in memory in order to use any of the VOICE MASTER commands. If "VM2.1N" or one of the others has not been loaded in, attempting to LIST or RUN the program will result in a syntax error when a VOICE

MASTER command is encountered. (But you can type in, save and load BASIC programs with VOICE MASTER commands without having a VOICE MASTER program in main memory.)

Now type the keyboard command LEARN 1. Have the microphone ready when you press the RETURN. Say a few words. The screen goes blank, which indicates that the Voice Master is recording. When you stop talking, the screen comes back on. This will happen if you pause too long in the middle of your talk, so don't pause if you want to get everything in. A buffer is used with recording which limits duration to about 8 seconds. While you are getting familiar with the system, do your talking in a quiet environment. Note: Don't be upset if the screen blanks before you say anything. The computer is constantly testing audio input to determine if anything of substance has happened. Each few tens of milliseconds, the computer decides whether or not something has occurred and, if not, it resets. But the screen remains blank if some stray sound starts the recording process.

Now type SPEAK 1 and press the RETURN key. If the volume on your TV or video monitor (with audio) is high enough, the computer will reproduce what you recorded. You can SPEAK 1 over and over again because it remains in memory. You can re-record by typing LEARN 1 over and over again, with either longer or shorter utterances. Or you can record as many as 64 different words or phrases (numbered 0 to 63) by typing LEARN 12, LEARN 5, LEARN 0, etc. The order doesn't matter. You can hear whatever you put in with SPEAK 12, and so forth. If you SPEAK a number that was not LEARN(ed), you will get a short tone "beep". During playback of a sound, the space bar or "Q" key resets playback to the start of the word.

The commands can also be numbered statements, and numbers representing individual words or phrases can be computed. This means that a variety of simple programs can be written for recording, with playback in any order. In order to erase your vocabulary from memory, type CLEAR n. The number n is a value between 16 and 160. This number specifies the "page" in memory where the speech that is next recorded will begin. Leaving off the number on CLEAR gives the default value of 64 (or what was previously specified on disk if a stored vocabulary is next loaded into main memory from disk). The default value is satisfactory for most programs. The CLEAR command removes speech but does not affect any programs in BASIC or the machine language program. (NEW removes only the programs that are written in BASIC, but will not erase speech in memory or machine language programs.)

The following program allows you to enter 12 words under computer control, with screen prompt. You may wish to CLEAR before starting any recording session just to be sure that leftover words from some previous session are removed. You may also want to NEW so as to remove any statements left over from some previous BASIC program.

```
10 FOR J=1 TO 12
20 PRINT "SPEAK WORD NO.";J
30 LEARN J
40 NEXT J
50 END
```

When you RUN the program, start entering words by voice. Just count if you can't think of something else.

You can hear each word after recording it if you insert the statement: 35 SPEAK J. (Insert the statement when in the READY mode and then RUN again.) If you don't like one of the words, get into the READY mode with RUN/STOP and replace it using the keyboard LEARN command (with the appropriate index number). This won't affect any of the other words.

The "END" statement in the above program is sometimes optional, as it is in this particular case. But sometimes it is required if a choice situation within a program calls for terminating the program, or if subroutines follow the main program. Routinely including an "END" statement can be justified as constituting a good habit.

Now erase the BASIC program with NEW, but do not CLEAR. The 12 words remain in memory. You can use them in some other program, in any order. Here is one with a DATA statement where the words representing the individual reference numbers are individually spoken.

```
10 RESTORE
20 READ K
30 SPEAK K
40 GOTO 20
50 DATA 5,4,10,1
60 END
```

where an "out of data" statement will appear when done. (In the above, the "END" statement can come before the "DATA" statement; "DATA" statements can be anywhere.) If words come back too close together in time, insert a simple software delay loop: 25 FOR J = 1 TO 1000: NEXT. Or Use the PAUSE n command where n is a number between 1 and 255

that specifies the number of 1/4 second increments in the pause.

SPEED is a statement (or command) that is fun to work with. First record a word or phrase with LEARN. Then change playback rate with a keyboard command SPEED 4. (SPEED 6 is the normal, or "default" value.) Now SPEAK the word and you get a slowed reproduction. You can combine different SPEED values in a single BASIC program.

And there is more. RATE changes the digitizing sample rate on recording, with a range of values from under 5000 to over 12,000 bits per second. Higher bit rates tend to reproduce speech with less distortion, but at the cost of greater memory requirements. VOLUME changes the loudness of the speech, which is recorded with up to 16 amplitude levels (including zero). This allows for interesting special effects such as an echo. On recording, the video screen always blanks. On playback you can have the screen on or off with SCREEN n.

A complete vocabulary can be saved on disk or tape, and loaded back into memory in the command mode or under program control. Except for non-standard disks and/or power line frequencies other than 60 hertz, a special fast vocabulary loading routine is employed (with "VM2.1" only) which is several times faster than the normal method used with the Commodore 64. Each vocabulary is saved on disk or cassette as a named program file. There is really no limit to the size of the vocabulary that can be put into a single BASIC program.

An interesting special use of RATE and SPEED achieves a substantial compaction of data so that less memory is needed for storing a vocabulary. Speech quality can be improved without corresponding decrease in memory requirements as well. When you record a word, talk rapidly so that the duration of the word is short. Also talk in a somewhat elevated voice pitch. When you play back the word, reduce the SPEED so as to reduce voice pitch to the original value or somewhat less. You may discover that the reproduced word is even more pleasing than the original, especially if the RATE used in the recording process is somewhat higher than normal.

Note: The VOICE MASTER disk is write protected. This means that you cannot change anything on the disk or store anything new on it. If you want to save a vocabulary or a BASIC program, you must first make a complete backup copy of the VOICE MASTER disk with the copy not being write protected (using the utility routine that may be provided

with the Commodore 1541 disk drive), or you must format a new disk and save specific files and programs on it. A special backup program on the VOICE MASTER disk allows you to make copies of "VM2.1" and other machine language programs. A later section in this manual discusses copies and backups in more detail.

SOME CLARIFYING COMMENTS

Words are numbered 0 to 63. If you don't like a particular word, then simply re-record it as a keyboard command. Or use a short one line program with a line number higher than anything else in your program and RUN this program at that line number. The short program addition can be removed subsequently. Memory space is automatically increased or decreased to accommodate the re-recorded word. If you attempt to SPEAK a word or phrase that was not previously LEARN(ed), you will hear a telltale "beep".

A word index number for SPEAK can be computed. An example is a talking keyboard (for the blind?). Press, for example, the "W" key and have the computer respond with the spoken letter W, or perhaps the phonetic alphabet equivalent "whiskey" familiar to airplane pilots. The vocabulary for the letters could be taken (arbitrarily) as word numbers 20 to 46: 20 for "A", 21 for "B",...up to 46 for "Z". The "W" key is changed to the decimal number ASCII equivalent, which is 87. (See Commodore 64 User's Guide.) If the number 44 is subtracted from 87, the result is 43, which thus specifies the letter "W". Note: We must subtract a number so that the VOICE MASTER limit of 63 is not exceeded. The same number 44 applies for every letter in this example. (The ASCII number for "A" is 64 and $64-44=20$.) A different number would be used if our index list started at a number other than 20.

A sample BASIC program which gets the proper input key and speaks the correct pre-recorded word is

```
50 GET A$: IF A$="" THEN 50
60 N = ASC(A$)
70 SPEAK N-44
80 GOTO 50
90 END
```

Of course, limits could be placed on the value of A so that only letters would respond, or double limits to include lower case letters as well as capitals. And, of course, some action, such as printing, might be carried out

before returning for another letter in statement number 80.

Another technique for selecting words is to separate numbers or string variables into elements. Suppose, for example, a program computes \$24.95 and you want the computer to say "The price is twenty four dollars and ninety five cents". This will require 8 words and phrases from your vocabulary. We will combine the words "twenty" and "four" to produce 24 rather than saying the actual number 24. This reduces memory requirements without compromising generality in any number talking program, whether it be a general purpose reporting cash register or an automatic price quotation service, perhaps responding to a touch tone telephone or a button on a display case.

We first SPEAK the starting phrase "The price is". Then we decide if we have one or more dollars. If the price is only 80 cents, we don't want to say "dollars". This requires a decision with slightly different pathways being followed. We look at the numbers to the left of the decimal point. We determine that the digit in the tens place is 2. So we say "twenty". In the units place we find 4, so we say "four". Then we say "dollars". Had the number been "twenty" instead of "twenty four", we would have gone direct to "dollars". Then if we have no cents, we stop. Otherwise we say "and" followed by the number, and finally we say "cents". We may want delay loops or use the PAUSE statement in order to avoid words coming too close together in time. We will get out words from SPEAK N statements where N can be determined with a series of IF-THEN or other conditional statements.

In LEARN(ing) a word, the VOICE MASTER screen immediately goes blank upon pressing RETURN. However, actual recording does not commence until an adequately long speech sample (or background noise) is received. If the speech sample is too short, the program presumes that no speech has in fact occurred and resets so as to conserve memory. (The critical time period is adjustable as described in the section on memory locations.) Once the program decides that a valid speech sample has occurred, and following the end of the speech sample, the screen will remain blank until an adequately long period of silence has occurred, or when the buffer is filled, or if recording is forcibly terminated by striking the "Z" key. Normally, the recording process automatically stops after a short period when no speech is detected. In producing a string of words for a given reference phrase, be sure to not pause too long between words else the computer will stop recording. It also stops recording when the buffer is filled, which takes about 8 seconds. The Voice Master records speech with 16

levels of intensity (0 to 15). Below level number 1, no sounds are recorded. At level 15 and above, the indicator light glows. Some peak limiting is acceptable, even desirable, in most cases. Dynamic range is then maximized so that both weak and strong sounds are retained in their proper relationship. But avoid excessive clipping else speech can become distorted with excessive background noise. Too much background noise can prevent a recording from stopping until the buffer has been filled.

You may occasionally wish to stop the recording process. This might be desirable in a noisy environment when the noise itself prevents cessation of recording until the buffer becomes filled. In such situations, noise may also force recording to begin prematurely. In this event, useful information should be given to the computer as soon as the RETURN key is struck. Pressing the "Z" key will force a halt to recording.

SAVING AND LOADING VOCABULARIES

Both your vocabulary and your BASIC program can be saved from computer main memory to disk or tape. The BASIC program can be saved the same as any other BASIC program with the SAVE command. If not already saved, your vocabulary can be saved with the command

PUT "filename",8

However, you will not be able to save anything on the master disk supplied with VOICE MASTER because it is write protected. You can run demonstration programs because memory for vocabularies, templates, and user programs is supplied by computer RAM (or from pre-recorded words). But you cannot save these programs unless you use a different disk. We suggest that you adopt a generally prudent policy of making a backup copy of the master disk, and then do all of your subsequent manipulation with this disk. (See later section on making a backup.)

"VM2.1" or comparable VOICE MASTER program must be present in main memory in order to save a vocabulary. Besides saving the vocabulary itself, you will also be saving the speed, rate, screen, and volume settings which were used in creating the vocabulary just prior to saving it. In addition, you will save the page number where speech storage begins as determined by the CLEAR n command.

A vocabulary is loaded from disk as

FIND "filename",8

---- OR ----
FFIND "filename",8

There are two FIND commands. One of these, FFIND, is several times faster than the other. The normal FIND speed is consistent with the standard Commodore 64 disk drive system. If you are not using the Commodore Model 1541 disk, and/or your power line frequency is not 60 hertz, then you may be limited to using only the standard speed FIND. (Vocabularies stored on tape cassette cannot use the fast routine either.)

If you have two or more different named vocabularies on your disk, and you wish to replace one vocabulary with another in the same BASIC program, then just FIND (or FFIND) the one you want. The old one is overwritten because a CLEAR is incorporated in FIND. You can make this swap automatically in a BASIC program because FIND can be a numbered BASIC statement. The start of speech when placed in memory from disk or cassette will be at the address originally used in recording the speech. You can add to or change the vocabulary after it is loaded, but the address of the start of the vocabulary remains unchanged. Thus you have access to an almost unlimited number of words and phrases! (Also recall that the fast loading routine, FFIND, is contained only in "VM2.1".)

Remember that you must have "VM2.1" or "VM2.1N" or "VM2.1R" in computer main memory if you want to record and/or play back words in a vocabulary. It will normally remain there while you program in BASIC until the computer is turned off (or if you have a power failure). Be sure to load your BASIC program after loading "VM2.1" or similar program. Otherwise, when RUN, the VOICE MASTER statements in your program will not be recognized and a syntax error will occur. As will later be described, a special procedure allows "VM2.1" to be loaded from a BASIC statement without destroying the BASIC program. "DEMO" and some other BASIC programs that implement speech recording and playback utilize this method.

SPEED, SAMPLING RATE, LEVEL, AND SCREEN

Words may be produced at different rates like a slowed or speeded tape recording. On a scale of 11, normal speed is 6 with 0 the slowest and 10 the fastest. Just before the SPEAK N statement, you can use a statement SPEED M to changed from the normal (default) value of 6. Because the M value can be computed, speed can be changed within a program. Once you have changed to a non-default value, you must insert SPEED 6 in order to get back to normal. The following example repeats a word eleven times, from slowest to fastest:

```
10 FOR J= 0 TO 10
20 SPEED J
30 SPEAK N
40 NEXT J
50 END
```

The sampling rate at which the speech is digitized can be varied using the command RATE N, where N is in the range of 0 to 10. A high sampling rate can improve speech quality but will require more memory. Normal sample rate setting is RATE 6 (the default value), which corresponds to 7100 bits per second. When you change the sample rate, it will be necessary to change SPEED as well in order to maintain natural voice pitch and rate of word production. That is, if you sample faster, you should play back faster. SPEED and RATE reference numbers are keyed together so that playback with a natural voice pitch level occurs when both numbers are the same.

There are 16 amplitude levels available, including zero. Speech uses all of these while recording. But loudness can be changed up and down with another statement. The VOLUME command has 16 values from 0 to 15 with 15 being the loudest. The normal (default) value is 15. Any setting less than 15 defines the maximum volume that speech will ever reach. Thus, a setting of 4 means that anything that was recorded louder than 4 is limited to 4. Anything less than 4 remains unchanged. The overall effect produces quieter speech but also results in a reduced dynamic range and increased distortion. An example of an echo effect is had by repeating a word 5 times at increasingly reduced volume levels with the program

```
10 FOR J = 15 TO 0 STEP -3
20 VOLUME J
30 SPEAK N
40 NEXT J
50 END
```

Yet another statement is SCREEN. During recording, the screen is always blanked. During normal playback, the display screen is normally not blanked. The screen can be forced to blank during playback with the statement SCREEN n where n is any positive integer. The statement SCREEN without a number returns to the default (unblanked) state. Blanking the screen on playback can improve speech quality, especially if recording has been done at a high RATE.

THE CLEAR COMMAND

The command is CLEAR n where n is a value between 16 and 160. Any other value will produce "syntax error". The CLEAR command erases all speech in random access memory (RAM). The number following the CLEAR statement tells the VOICE MASTER the point in memory, or "page", to start storing the new speech you are about to record from the microphone.

For example, the command, CLEAR 64, will erase all previously stored speech and will start storing newly introduced speech beginning at page 64 in memory. A page of memory is defined as 256 bytes. Therefore, page 64 means that speech starts at location $64 \times 256 = 16384$ (decimal). The lower the page number you define, the more memory available for speech, but then there is less for your BASIC programs. The command, CLEAR 160, will start storing speech "behind" the BASIC ROM, giving the full amount of memory available for BASIC programs, but only 6000 bytes (or about 6 seconds) for speech. Recall that, in the Commodore 64, the ROM used to hold BASIC can be swapped with RAM at the same address when working with assembly language. Storing speech "behind" the BASIC ROM means that storage is at the same addresses used by the BASIC ROM and hence will not interfere with the longest possible BASIC program.

You can find out how much memory you have available for your programs in BASIC by typing

```
PRINT FRE(0)-(FRE(0)<0)*65536
```

The default condition is CLEAR 64, which occurs when you first load in one of the VOICE MASTER machine language programs. Typing CLEAR without a number will set the memory to whatever was set previously, or to the default value if no other number has been specified. If a non-default value has been set in and you want to change this, then the desired number must be written down explicitly. The current memory page setting is stored in

zero page memory location 56 and can be determined with the keyboard command PRINT PEEK(56).

The functioning of CLEAR when loading a vocabulary from disk or tape is different in some respects from that described above. A FIND or FFIND always CLEAR(s) and the stored vocabulary contains the starting page number that was used in producing the original vocabulary that was PUT. Whatever number might have been used with a CLEAR prior to FIND(ing) the vocabulary is simply ignored. It is not easy to change the starting page number stored on disk or cassette memory.

THE IMPORTANT TASK OF MAKING A BACKUP

VOICE MASTER software contains a program that will make copies of VOICE MASTER machine language programs (only). Make sure that you have a formatted diskette ready before using. To operate, type

LOAD "BACKUP",8

and type RUN. "BACKUP" is menu driven and all necessary instructions are presented on the video display. VOICE MASTER programs written in BASIC can be saved in the normal way. That is, load the program from one disk to main memory, and save it on the backup disk.

There is a Commodore 64 backup program that will transfer information from the write protected master disk to a properly formatted disk. The formatted disk will not be write protected so that you can scratch files, save new material, etc.

Because the Commodore disk drive operator's manual is not particularly easy to read, we give a brief and specific explanation on how to proceed. First, to format

OPEN15,8,15
PRINT#15"NEW:your name,your id number"

where the name is an acceptable file name and the id number is any 2 characters. (There is a shorthand way to format which involves just one statement. See Commodore disk drive manual.) After the format is complete, remember to CLOSE 15. File number 15 is the command channel file. If this file is used with other files, close the other files first, then CLOSE 15.

To make a backup copy, insert the utility disk supplied with the Model 1541 disk drive and type

LOAD"SD.BACKUP.64",8

exactly as shown. Then RUN. A menu appears and you will be told exactly what to do. The source disk is the VOICE MASTER disk. The destination disk is your new formatted disk. Unfortunately, the Commodore manual is somewhat misleading and incomplete. Although not specifically stated, it is not possible with the software provided to copy individual machine language programs from one disk to another. BASIC programs are copied by first loading into main memory and then saving to the new disk. VOICE MASTER vocabularies (and template sets in recognition) are saved to a disk with PUT (TPUT for templates) and require that a suitable VOICE MASTER program reside in main memory.

Note: Recently manufactured Model 1541 disk drives appear to come with the program cited for use here, "SD.BACKUP.64". Earlier versions may not have this program, but "UNI-COPY" or "COPY/ALL" instead. These will not work on a one-disk system--they copy from one disk drive to another. The latest model disk drive, Commodore Model 1571, has the above backup program as well as "SD.COPY.64" which will copy individual programs and may be an alternative to "BACKUP" on the VOICE MASTER disk.

LOADING VOICEMASTER FROM BASIC

Earlier in this manual, we discussed how to load the essential machine language program using keyboard commands. That is, LOAD"VM2.1" or a comparable program, SYS to 49152, and then NEW. When you NEW, you automatically reset pointers in page zero to the BASIC program you are about to write or load. If you have already written your BASIC program, a NEW will promptly erase it. Clearly, you don't want to NEW after a BASIC program has been entered, but if you do not NEW, pointers to some subsequently written program will not be correctly set and you will not be able to add BASIC statements. If you try, you will get an "out of memory" error.

So what can we do? If a statement in a BASIC program serves to load in "VM2.1" (or "VM2.1N", etc.), then the pointers are initially set with the start of the BASIC program without requiring NEW. The method described here is the one that you will probably use in a practical program where only a single BASIC LOAD command is needed to get everything to RUN. The procedure is, in effect, to LOAD"VM2.1",8,1 and then SYS 49152 so that the machine language program begins running at location 49152 (C000 Hex). But there is one "peculiarity" in Commodore 64

BASIC. The necessary loading statements must be written as

```
10 IF A=0 THEN A=1:LOAD"VM2.1",8,1
20 SYS 49152
```

where it is assumed that these are the leading statements in the BASIC program. The IF-THEN statement is necessary. Without this, a LOAD would initialize and then LOAD again in an endless loop. With the IF-THEN statement, the condition is not met the second time through statement number 10 so that the statement after the colon is not reached and the program proceeds directly to 20. Loading in this way from a BASIC program is the preferred method after you have completed the development and debugging of your BASIC program and are ready to get a vocabulary into main memory as well. It is instructive to LIST some of the programs in the main "MENU" for study. But you will have to RUN "MENU" before you can LIST any of them so that VOICE MASTER statements will be recognized (i.e., running "MENU" loads in "VM2.1R").

What happens if your BASIC program has a group of instructions before the pair of instructions in the example above which does the loading? Subsequent to the load, the program returns to the beginning of the BASIC program and thus this initial group of instructions will be carried out twice. This is characteristic of BASIC and exists in order to chain long programs drawn one after the other from disk memory. The "IF A=0 THEN A=1" trick prevents a reload of "VM2.1", but does not prevent going over preceding BASIC statements. Thus, put the load instruction at the start of your program. An alternative is to RUN starting at the loading instruction (which must in any event occur before any VOICE MASTER commands).

THE PLAY/FIND PROGRAM

"PLAY/FIND" is a simplified machine language version of "VM2.1N" which contains the playback and vocabulary loading routines only. It does not include the recording function. This program is for those who want to use speech response in their own programs from pre-recorded vocabularies. As in the case of loading the machine language program "VM2.1N", it will not be possible to change the BASIC program after loading "PLAY/FIND". It is presumed that you have already written and checked your final program and next want only to get a vocabulary into main memory. When you get to the point in a BASIC program where the assembly language program is to be installed, then insert the suitably numbered statement

```
150 IF A=0 THEN A=1:LOAD"PLAY/FIND",8,1
```

where this loading must precede the loading of a vocabulary. Loading the vocabulary is done with

```
220 SYS 49155"filename",8
```

Note in the above that "PLAY/FIND" has been loaded without a following SYS to indicate where the program is to begin running. Such a SYS was used in loading "VM2.1". The reason for the difference is that "PLAY/FIND" does not utilize wedged-in BASIC commands. If you wish to load the program from the keyboard (not from a BASIC program), then type LOAD"PLAY/FIND",8,1 (return) and then NEW.

"PLAY/FIND" has been simplified in part by removing the wedged in BASIC commands. These must all be replaced with POKE(s) and SYS. The SYS in the above example replaces the BASIC FIND command. Although the programming is somewhat more complicated, there are advantages in reduced memory storage requirements. It is important to note that "PLAY/FIND" does not use the fast vocabulary loading method.

As stated above, special considerations are needed when using "PLAY/FIND". Because the recording part is deleted, some of the simplified BASIC statements must be modified. Statements such as SPEAK, VOLUME, and others must be replaced with POKE and SYS statements. Just what numbers are to be POKE(d) can be determined from memory locations given later in this manual. For example, SPEAK 0 is replaced with POKE 151,0:SYS 49152. Location 151 contains the reference number for the word to be spoken, and location 49152 is where the machine language program begins.

The following program serves to illustrate some of the above considerations. This BASIC program loads "PLAY/FIND", FIND(s) "GREETING" (also used in the "DEMO" program but otherwise is a separately named vocabulary on disk), and speaks the first two phrases in "GREETING" with the second phrase echoed.

```
10 IF A=0 THEN A=2:LOAD"PLAY/FIND",8,1
20 SYS 49155"GREETING",8
30 POKE 151,0:SYS49152:REM SPEAK 0
40 FOR J=0 TO 15 STEP 3
50 POKE 253,J:REM VOLUME 15-J
60 POKE 151,1:SYS49152:REM SPEAK 1
70 NEXT J
80 POKE 253,0:REM RESTORE MAX VOLUME
```


In this example, the contents of memory location 151 will reference the phrase number to be spoken, and the contents of location 253 are subtracted from 15 to give a volume setting. Note that line 50 is different to what you would use with the VOLUME command when running "VM2.1". This is because the value in location 253 is not the actual volume setting, but it is a value to be subtracted from 15. For example, the BASIC command POKE 253,8 is equivalent to the VOICE MASTER command VOLUME 7.

There is an alternative to using "PLAY/FIND". That is, simply use "VM2.1" or "VM2.1N". If your only interest is in playing back words from a pre-recorded vocabulary, then you will be carrying along extra memory to facilitate VOICE MASTER commands (and fast loading parts of the machine language program if you use "VM2.1"). This may not be a problem in many practical applications, but might be objectionable in writing software that is to be produced in quantity. "PLAY/FIND" is efficient and short compared to "VM2.1".

OTHER PROGRAMS AND COMMENTS

When running a BASIC program, you can change the playback characteristics at any time by stopping the program with the RUN/STOP key. Then you can change speed, volume, screen, etc. using VOICE MASTER commands for "VM2.1" or the equivalent POKE(s) for "PLAY/FIND". Then continue the program with CONT.

There is a program "peculiarity" with Commodore BASIC. All normal BASIC commands work with the new commands created for Voice Master. There is a minor exception which also happens to occur with the extended version of BASIC when installed in the VIC 20. That is, the IF-THEN, ON-GOTO, and ON-GOSUB should have a colon inserted where none would ordinarily appear. The following examples illustrate this.

```
IF W=3 THEN: SPEAK J
ON X: GO TO 500
ON X: GOSUB 550
```

The first example above could be written with an indirect SPEAK command without having to use the colon as: IF W=3 THEN GOTO 50, where 50 is the line number containing the SPEAK J statement.

SPECIAL DATA CHECKING PROGRAM

This is a sample program that you can type in and use in your future work. The purpose is to read off numbers previously typed into data statements so that you can check for accuracy. As you run your finger along the printed listing of the program, Voice Master speaks the numbers. If there is an error, then simply note the error for later correction and continue. You can hit the RUN/STOP key if you want to stop or rest at any time. Then type CONT to proceed. There is a limitation in that you cannot change data statements without resetting to the start of the data.

You must have three programs in the machine at the same time, along with the necessary vocabulary. You can create the vocabulary at any time. We presume you do this for the digits 0 through 9 using a program such as

```
10 FOR J=0 TO 9
20 LEARN J
30 NEXT
40 PUT"NUMBERS",8
50 END
```

where it is presumed that you previously have loaded "VM2.1" (or "VM2.1N", etc.). Note that the vocabulary is named and saved in the above program.

The three programs you will need are the program with the data statements that you wish to check, the checking program itself, and one of the VOICE MASTER programs, (such as "VM2.1" or "PLAY/FIND"). Of course, do the loading so that you don't destroy BASIC programs.

As an example, suppose you have data statements for checking as numbers 490 and 500. Put a special number as a final data statement so that the checking program will stop automatically when all data have been checked. The program being checked is presumed to END at statement number 1000. Have your checking program start at an even higher number and then RUN from this number. The sample checking program is as follows:

```
.....
490 DATA 0,303,10,30,99,102,200
500 DATA 1,34,67,128,23,0,20
501 DATA -999
.....
1500 SCREEN 9
1510 FIND"NUMBERS",8
1520 READ J:IF J<-900 THEN END
```

```
1530 PRINT J
1540 A$=STR$(J)
1550 FOR K=2 TO LEN(A$)
1560 B$=MID$(A$,K,1)
1570 B=VAL(B$)
1580 SPEAK B
1590 NEXT K
1600 PAUSE 4
1610 GO TO 1520
1620 END
```

You could elaborate on this program to permit editing during operation when you detect an error. This requires keeping track of the number of data statements that have been checked (by counting passages through the READ statement). After the error is corrected, the program is run again from statement 1520 with a subroutine inserted so as to quickly scan through all the data statements up to the one following the correction.

BLANK

PART 2: SPEECH RECOGNITION

PRELIMINARY COMMENTS AND SUGGESTIONS

Some of the terms used in this section will not be explained until subsequent sections. What is presented here will nevertheless be useful as you learn how to work with voice recognition. The demonstration program "RECOG-DEMO" will also later be described in some detail because of its general ability to illustrate the several features of VOICE MASTER.

Be sure that you are properly calibrated!

Place the headset on your head with the earphone on your left and the other end of the wire behind your right ear. Adjust the microphone element so that it is within one inch and directly in front of the lips, perhaps a little towards the nose so as to assure pickup of nasal sounds. The microphone permits a wide range of position adjustments to be made. Adjust the "GAIN" control so that the indicator light on VOICE MASTER glows occasionally when you say "ah". In a noisy environment, it may be necessary to lower the input level and speak loudly. But in most cases, once you have set the level with "GAIN", you will not need to re-adjust. You may wish to review the section, "CALIBRATION CHECK", especially on setting GAIN with the "BAR" program. It is very important that GAIN be set correctly so that speech will have zero level before and after a word; it is this cue that tells the computer when a word starts and when it stops.

VOICE MASTER implements speaker dependent, isolated word recognition. It is meant to work best with the person who originally TRAIN(ed) the system. Good results with different voices can be achieved if words are multi syllable, sound quite different, and the vocabulary is small. You can adjust the threshold values to make the system less sensitive to different speakers, but possibly with an increase in the number of substitution errors.

Don't shout or try to talk in an overly precise way. The recognition program works best if the words are produced in a normal and relaxed manner. If you use VOICE MASTER in a stressful environment where it may be necessary to shout or speak loudly, then make sure that you have TRAIN(ed) the system to the way that you will probably speak.

The template for a word to which an unknown word is matched will often be found to be a more reliable indicator

if it is the average of two TRAIN(ings). Averages over three or more TRAIN(ings) can also be used, but two seems to be a good value; averaging over more can in some cases reduce accuracy and reliability. When multiple TRAIN(ings) are desired, it is suggested that the same word not be TRAIN(ed) sequentially. For example, in TRAIN(ing) for the ten digits, go through the entire number list once, and then do it a second time. When a person says a word twice in a row, the average value may not be a good indicator or the average when the word appears more randomly in time.

Be aware of extraneous noises just before and just after speaking, such as lip smacks, tongue and teeth clicks, and breath noises. These sounds are acoustically similar to plosive bursts and/or fricatives and could be mistaken for such. Noises at the start and/or end of a word can be especially troublesome, in part because they confuse the identification of just when the word is supposed to start and end. (A plosive is a brief burst of sound as in the letters "t" and "p".)

Be aware of how you release final plosives like "t" and "p". It is often optional in ordinary speech to release the final plosive. Consider, for example, the final "t" in the word "eight". You may not even pronounce the "t", or replace it with a weak "tah". Or you might put the "t" close to the end of the vowel part of the word eight, or wait awhile before doing this. Problems can be especially severe when you try to differentiate between words such as "ache" and "ate" (or "eight") and "ape". Use of such similar sounding words as these in a single vocabulary is not considered to be good practice; even a human listener will have trouble. In general, try to avoid using rhyming words and words with final plosives. Such words are similar with respect to both the vowel part and the final plosive.

Generally, the larger the vocabulary and the more similar sounding are the words, the larger will be the error rate. If accuracy is a problem, make use of the 8 word sub-groups in the RECOG command.

Multi syllable and acoustically different words give the best results. Consider the telephone operator who pronounces the number "five" in a rather special way so as to differentiate it from "nine", or the pilot who says "zero" instead of "oh", and "niner" for "nine". If humans have such a hard time with similar sounding words, is it really fair to demand that the computer do much better?

SPECIAL KEYS AND SYMBOLS

When the computer is waiting for a word to be input, a question mark appears in the upper right corner of the display. If a stable unchanging "?" does not appear when you do not speak, then either your VOICE MASTER needs calibration or there is excessive background noise. When you talk, the "?" becomes a different active symbol; and both disappear when you finish.

If it is desired to put active recognition on "hold", press the up arrow on the keyboard (not the cursor arrows). An up arrow will replace the question mark. In order to return to the recognition phase, press the up arrow key again. Terminate the speech recognition function with The RUN/STOP key.

Finally, pressing the "Q" key will put the number 253 in memory location 151 (in zero page) which gives the programmer a kind of "interrupt" capability if location 151 is being scanned.

RECOGNITION PROGRAMS

Three machine language programs incorporating speech recognition are on the VOICE MASTER disk. All have the same objective of constructing what we will call "templates" for words which are to be recognized. There is one template per word. A template is a collection of "cues" representing certain attributes of the word. These are retained in memory for comparison with a template that is constructed from an incoming unknown word to be recognized. The best match defines the word that is recognized. It is presumed that word recognition will be combined with a user program.

One of the three programs, "VM2.1R", is intended for use with speech playback so that the computer can both talk and listen in the same user program. The wedged-in BASIC commands for playback and recognition must be combined because the Commodore 64 is not able to handle more than one set of wedges at a time. The second program is for those who want only the word recognition feature, but still desire the simplicity of wedged BASIC commands: "VR1.0". This program is considerably shorter than when speech playback is included as well. The third program, "VOICE/RECOG", is also for speech recognition without playback. But in this case, POKE(s) and SYS are used in place of wedges, similarly to the way that "PLAY/FIND" is manipulated.

Programs with wedges are loaded in the same way as described for recording and playback:

```
LOAD"VR1.0",8,1
SYS 49152
NEW
```

"VR2.1R" is loaded the same way. If loading is within a BASIC program:

```
10 IF A=0 THEN A=1:LOAD"VR1.0",8,1
20 SYS 49152
```

For the program without wedges:

```
LOAD"VOICE/RECOG",8,1
NEW
```

which is loaded from a BASIC statement as

```
10 IF A=0 THEN A=1:LOAD"VOICE/RECOG",8,1
```

The program "VR1.0" is considerably shorter than "VM2.1R", which is the merged combination of "VM2.1N" and "VR1.0". The recognition program "VR1.0" is located "behind" the kernal ROM. Except for saving and finding templates, it makes no use of the kernal (whereas "VM2.1" and "VM2.1R" make frequent use of it).

The third program, "VOICE/RECOG", does not utilize a wedge. Commands must be entered using POKE(s) and SYS. The recognition program without wedges can be combined with "PLAY/FIND" for an all-machine-language speech input-output program.

PRINCIPAL STATEMENTS

Assume that you have put "VR1.0" (or "VM2.1R") into computer main memory. The next task is to create word templates so that word recognition will be possible. Use the keyboard command (or BASIC statement)

```
TRAIN n      (n in the range 0-31)
```

where the number identifies the particular template, similarly to the way the number identifies a word in a vocabulary for speech playback. The numbers can be produced in any order and not all need be employed. It is suggested that TRAIN(ing) numbers be put into groups as

0-7, 8-15, 16-23, and 24-31. Comparing templates for recognition can address one or more of these groups selectively.

If the same word and index *n* is TRAIN(ed) twice, the template will contain the average (a special kind of average) of the cues for the word as will be used in recognition. An average can be acquired for any number of repetitions of the same word with the same index. We suggest that each word be TRAIN(ed) twice. Once will usually not be enough. Three or more times may not provide much of an advantage and could possibly lead to a degradation in performance. (Averaging is carried out using dynamic time warping.)

If you want to erase a template so as to replace it with a different word, or try again to get an improved template, then use the keyboard command (or BASIC statement)

BLANK *n* (*n* in the range 0-31)

where *n* designates the particular template to be erased. BLANK without a number erases all templates. This command is analogous to CLEAR for speech recording and should be used with considerable caution.

Once the desired number of templates has been created, we are ready to recognize. With the microphone held and used properly, type the keyboard command or BASIC statement

RECOG *n* (*n* in the range 0-4)

where RECOG without a following number is equivalent to RECOG 0. (The number *n* will be described shortly.) Then push the RETURN key and say the word. The program waits for speech input the same as when recording. External noise must be avoided (but short bursts will be ignored the same as in recording). After the word to be recognized has been entered, the program scans the selected templates. If a suitable match is found, the index number of the template will be placed in memory location 151 (decimal) where it can be PEEK(ed) in order to be used in some desired manner. One way to indicate recognition is to simply PEEK the location and then print out the value in that location on the screen as PRINT PEEK(151).

A symbol on the screen helps to verify the recognition process. When a recognition is requested, a "?" appears at the upper right corner of the screen. When speech is actively being sampled, a randomly changing graphical

symbol appears in place of "?" Once a phrase has been correctly accepted (i.e., word length not too short), the symbol will disappear.

The RUN/STOP key is used to terminate TRAIN. The up arrow key (not cursor key) is used to put recognition on hold as, for example, when you wish to talk to someone or shuffle paper. The question mark is replaced with the up arrow. Pressing the up arrow a second time returns to the recognition mode. The "Q" key does a "pseudo" recognition with the number 253 being placed in location 151.

PARTITIONING

It takes time to scan over templates. The maximum of 32 templates for RECOG 0 requires more time for making comparisons than when fewer templates are scanned. But even with 32, the time required is considerably less than one second. The range of scanning can be limited to only certain groups of template numbers as:

RECOG or RECOG 0 scans all 32
RECOG 1 scans 0-7
RECOG 2 scans 8-15
RECOG 3 scans 16-23
RECOG 4 scans 24-31

Two of the groups of 8 templates each can be scanned with a command such as RECOG 1,3 (in the order 1,3). Three are scanned with RECOG 1,3,4, and so on.

If there are, say, 6 words having templates in memory, then the smallest possible scan will cover two additional templates for a total of 8. In order to avoid accidental false recognition indications, the two reference numbers not having meaningful templates should be BLANK(ed). (Or BLANK everything before TRAIN(ing) the 6 words.)

By restricting scanning to subgroups of index numbers, response will be faster. Many practical situations involve only a few words. If the vocabulary is chosen so that words are quite different from one another, accuracy can be enhanced.

Reliability can be further enhanced with a two step recognition process and partitioning. The first recognition of only a few distinctly different words can then select the next group of 8 words so that no more than 8 templates need be scanned at a time. Suppose, for example, we have 2 very different sounding words in the

first group. One of these words can call up a second group, and the other can call up the third group. Then if there is a word in the second group that is very similar to a word in the third group, the chances of confusing these two words is small--essentially equal to the chance of confusing the two very different words in the first group. A possible application of this sort of segmenting could be the voice command of "pull down" menus. A "literal" example applies to a fast food restaurant where you first select categories such as "sandwich", "drinks", or "deserts". After the first selection, the next group of templates is selected for "cola", "root beer", "orange", and so on.

Vectoring through segmentation can also be used to bring new vocabularies of templates from disk. Because each template requires only 96 bytes, the loading process can be rapid. However, each template group on disk includes all 32 words, or 32×96 bytes.

SAVING AND LOADING TEMPLATES

The templates created by TRAIN(ing) can be saved on disk and acquired from disk within a BASIC program. The commands (statements) are

```
TPUT"filename",8  
TFIND"filename",8
```

Because templates can be named and saved, there is really no limit to the number of words that can be stored for subsequent recall and recognition by the person who prepared the original templates. Partitioning can be carried out almost without limit by using one recognition to load in a particular vocabulary of templates for subsequent recognitions.

Although recognition from templates is nominally restricted to words spoken by the person who originally made the templates, effective recognition can sometimes occur when the voice to be recognized is different. Some word sounds are more resistant than others to changes in talkers, and such special vocabularies can be useful in speaker independent applications. Experimentation by the user can best identify such opportunities.

THRESHOLDS

When a word is compared to a template, a numerical score is computed. The smaller the score, the closer is the match. The maximum nominal (default) score value is

300 (decimal). If no match is better than this, no recognition is indicated. The number 255 can be made to appear on the screen if such an error occurs with the statement PRINT PEEK(151).

Generally, the match with the lowest score (less than 300 for the nominal or default value) is selected. However, there is a differential threshold as well (default value of 30) which also results in no recognition. This situation applies when the difference between the best match (lowest score) and the next best choice is not large enough. The result is the number 254 in location 151. If two words have many similar properties and might be confused, a small number for the minimum differential score would be used, especially if a no-recognition condition is preferable to making a substitution error.

Threshold values can be changed from their default values with BASIC commands

```
MIN n      (n in range 0-100. Default 30)
MAX n      (n in range 0-700. Default 300)
```

Default values exist unless specifically changed by one of these commands. They will remain at selected values until returned to default values (or on reset and reload).

Minimum and maximum limits can have a significant effect on recognition accuracy. Optimum values may depend on the particular speaker as well as the particular vocabulary. Best results require some experimentation.

THE "VOICE/RECOG" PROGRAM

It will be recalled that the program "PLAY/FIND" is the speech recording and playback program without wedges and that BASIC commands are replaced with SYS commands. The purpose of having "PLAY/FIND" is to provide a program of minimum length which can be embedded in some user program. An analogous situation applies to "VOICE/RECOG" in that there are no wedges and it is meant for embedding in a user program.

Because "VOICE/RECOG" does not have wedges, commands such as TRAIN, etc., must be replaced with POKE(s) and SYS. Except for RECOG, these equivalences are:

```
TRAIN n      POKE 151,n: SYS 49920
BLANK        POKE 782,255:SYS 49932
BLANK n      POKE 782,n:SYS 49932
TPUT         SYS 49926"Filename",8
```

where n is the word index number for training or blanking. The special case for n = 0 (or no number at all) will cause all 32 templates to be BLANK(ed).

Minimum and maximum threshold values require POKE(s):

MIN n:	Low value to location	45607
	High " "	45608
MAX n:	Low value to location	45609
	High " "	45610

Examples:

If n in MAX n is 300, then
 POKE 45609,44
 POKE 45610,1

If n in MIN n is 30, then
 POKE 45607,30
 POKE 45608,0

An explanation for the values of POKE(d) numbers in the above examples may be helpful. The number 300 converts to high and low byte values as follows: 256 goes into 300 once with a remainder, thus the high byte contains the decimal number 1 for once. The remainder is 44 and this goes into the low byte. Similarly for the number 30: 256 does not go into 30 at all (thus zero in the high byte), and the remainder of 30 goes into the low byte.

Replacing the wedged in RECOG statement is a little more complex. On initial installation, the ruling location 49935 contains a null. The entire 32 templates will be scanned not once, but 4 times. This condition should be modified with a series of two to 4 lines of code. The last line must POKE in the number 255 so that scanning stops. The pertinent statements are:

RECOG n	POKE 49935,n
	POKE 49936,255:SYS 49923
RECOG 2,4	POKE 49935,2
	POKE 49936,4
	POKE 49937,255:SYS 49923

where n=0 (or no number at all) results in all 32 templates being scanned once. The second set of instructions illustrates how to handle recognition of 16 words using two specified groups of 8 templates each, with scanning in the order specified.

DEMONSTRATION RECOGNITION/PLAYBACK PROGRAM

The program "RECOG/DEMO" of interest here is selected from "MENU". It is discussed in some detail because it illustrates the use of VOICE MASTER in a composite program involving both recognition and audio response. If you want to RUN "RECOG/DEMO" without first loading and running "MENU", then LOAD "VM2.1R", SYS 49152, and then NEW in order to install the necessary "VOICE MASTER" program. Then you can LOAD "RECOG-DEMO", 8.

You are first asked to TRAIN the names of colors and a terminator "stop" so as to prepare templates. You go through this routine twice. Immediately following this, you are asked to produce words that will be stored for voice response to verify the recognition.

After TRAIN(ing) and LEARN(ing), the recognition mode commences. Say one of the words, but not until a "?" appears at the upper right corner of the screen. If the computer recognizes this word, then the color name and index number are printed on the screen and the color on the screen changes appropriately (if you are using a color monitor). Also, the name of the color is repeated back in your own voice. (Note: The index number is one less than the number of the named color as specified on the Commodore 64 keys. That is, black is number 1 on the key, but the index is 0.)

If recognition fails, the screen prompts with "say again" and the error code 254 or 255 is printed as well. This number signifies the type of error (minimum or maximum threshold, respectively).

Instead of colors, you can TRAIN for the numbers 1-9 and some suitable terminating command. Or any other words that you might choose. If you list the BASIC program, you will be able to determine how to add more words for recognition, get different screen printout, etc. This short BASIC program provides useful and meaningful instructions on using Voice Master for word recognition combined with voice response. (Note: In order to get a LIST(ing) of the program, you must have the VOICE MASTER program "VM2.1R" in memory. Otherwise, the computer will not recognize the wedged in commands and a syntax error will result. If "RECOG/DEMO" is loaded from "MENU", "VM2.1R" will be in memory and LIST will function after pressing RUN/STOP.)

The recognition phase can be terminated at any time with the last trained word, such as "stop", which returns

to the main "MENU" if "MENU" remains in memory, or to READY otherwise. Or hold down the RUN/STOP key, say something, and wait. The recognition phase can be restarted with the keyboard command CONT. After stopping, you can change threshold values with MIN and MAX, SPEED or VOLUME of response, etc., and then CONT in order to continue with modified parameters. Or you can re-TRAIN or re-LEARN individual indexed words with keyboard commands before CONT.

The program does TRAIN(ing) separately from LEARN(ing) This has the advantage of permitting response to be different from the recognized word if so desired. For example, response could be in a foreign language.

COMMAND SUMMARY (Recording/Playback)

SPEAK N. Designated word or phrase in the range 0 to 63. The number N can be computed. This statement (or command) plays back the designated word or phrase through the TV monitor or external amplifier. If you attempt to speak a phrase that does not exist, you will hear a beep. The space bar or "Q" key resets to the start of the word being produced.

LEARN N. Similar to SPEAK N except that it causes a word or phrase to be recorded in main memory, using the Voice Master digitizer with the microphone. Re-entering a phrase replaces the old one and memory space is adjusted accordingly. A buffer limits each speech period to about 8 seconds. Recording stops when a certain silent period occurs. Some pauses between words may cause an unintended stop in recording. Recording does not start until sound has been above zero level for a certain minimum time. Such things as clicks will thus not start the process. In a noisy background, start of recording may be spontaneous and recording may continue to the end of the buffer. Recording can be forcibly stopped with the "Z" key.

PUT"filename",8. This saves a vocabulary on disk, starting on page n (see CLEAR). For tape cassette, the comma and number 8 are deleted. Also saved are speed, volume, and screen settings.

FIND"filename",8. Recovers named vocabulary that was previously PUT. Retains same starting page address and speed, volume, and screen settings.

FFIND"filename",8. This is a fast loading version of FIND. Using it in place of FIND may require the Commodore Model 1541 disk and a power line frequency of 60 hertz. The fast FIND is not applicable to cassette storage or to the special machine language program "PLAY/FIND". FFINd is contained only in "VM2.1".

CLEAR or CLEAR n. Number n either omitted or given in the range 16-160. Using any number outside the allowed range gives "syntax error". Clears a vocabulary from main memory and prepares for introducing a new vocabulary from the microphone. Sets all Voice Master parameters to their normal (default) values. The number n specifies the page in memory where the vocabulary is to start. With n left unstated, the default value of 64 is inserted. Decimal memory location is $256*n$. Loading a vocabulary from disk or cassette automatically applies the correct page number as used in the original recording process and CLEAR is

automatic such that all previous CLEAR statements or commands are ignored. The number n applicable to a vocabulary on disk or cassette tape cannot be changed.

SPEED J. Changes response speed like a speeded or slowed recording. J is in the range 0 to 10 with 0 the slowest and 9 the fastest. Normal (default) value is 6.

VOLUME J. Changes output volume. There are 16 levels, 0 to 15, with 15 the loudest and 0 being zero. Amplitudes greater than 15 in the digitizer are limited to 15. Above 15 the LED limit light on the Voice Master glows and the rightmost bar in the program "BAR" turns red. (With a monochrome monitor, shading will change rather than color.)

PAUSE N. This acts like a software timing loop and produces a fixed delay. N is the number of one quarter second delay increments. Example: N=20 gives 5 seconds delay. For increments less than one fourth second, use a regular timing loop or just some dummy "time wasting" statement.

SCREEN or SCREEN n. During speech playback, the normal (default) condition is for the video screen to remain on. If the screen is to be blanked during playback, use the statement SCREEN n where n is any positive integer. To return the screen to the default condition, use SCREEN without a number. Quality of speech during playback will be somewhat better with the screen blanked, especially if recorded at a high RATE.

RATE J. Changes the sampling rate at which the speech is digitized. J is in the range of 0 to 10 with the normal (default) value 6. Details of rates are given under the section on memory locations. Note that a nonstandard RATE of recording will not play back at a normal speaking rate unless SPEED has the same index number.

IMPORTANT MEMORY LOCATIONS (Recording/Playback)

Many of the memory locations in this section refer to a "BASE" address, which is defined in the CLEAR statement. If uncertain, determine this address from the computer by evaluating the number in page zero location 56 as: PRINT PEEK(56)*256 (decimal).

Memory location 56 (\$38 hex). Contains the page number of memory where speech data begins.

Memory location 151 (\$97 hex). Current phrase number recorded or spoken. Range is 0 to 63.

Memory location 253 (\$FD hex). Volume setting to be subtracted from 15. Range is 0 to 15.

Memory locations BASE to BASE+255. These 256 bytes of memory define starting and ending addresses of where a phrase is stored in memory. For example, to find where the 7th phrase is stored, multiply 7 times 4, which is 28, to add to BASE. PEEK this memory location to yield the low order byte of the start of phrase 7, and the next location is the high order byte. The next two locations are the low and high order bytes, respectively, for the ending address for phrase 7. The following BASIC program finds starting and ending locations for any phrase (or word) number, with the specific example being for number 7.

```
10 WORD=7
20 PRINT"BASE=";PEEK(56)*256+(WORD*4)
30PRINT"START=";PEEK(BASE)+256*PEEK(BASE+1)
40 PRINT"END=";PEEK(BASE+2)+256*PEEK(BASE+3)
50 END
```

Note: This program (and the following one) will not work if CLEAR has used the numbers 159 or 160. The reason is that the leading page of information must not lie under BASIC ROM else needed data cannot be acquired.

Memory Locations BASE+256 and BASE+257. These two memory locations define the current top of speech memory. To find out how much memory you have used for your vocabulary, run the following program:

```
10 PRINT"BASE=";PEEK(56)*256
20 PRINT"USED=";PEEK(BASE+256)+256*(BASE+257))-BASE
```

Memory location BASE+259. Total number of recorded phrases. Range is 0 to 63.

Memory location BASE+260. This contains a number which determines the length of the shortest phrase that can be recorded. Normally set to 10. Poking a number smaller than this allows very short utterances, such as individual phonemes, to be recorded. But if it is set too low, then breathing or lip noises or short clicks may start the recording process and prevent it from ending at the end of the intended speech sound. This is not the same as background noise. Very short intervals can be utilized if recording is carefully done.

Memory location BASE+264. Screen flag. Setting to a non zero value will turn off the screen during speech playback.

Memory location BASE+265. Recording rate setting. This is not the same as the value specified in the RATE command, but is a setting used by the real time counter in the Commodore 64. A value of 128 corresponds to RATE 6. A higher number speeds up the recording sampling rate and a lower number slows down the sampling rate. Practical range for this is 16 to 200. The following table explains the relationship between the RATE setting, the value stored in BASE+265, and the approximate sample rate:

<u>RATE</u>	<u>BASE+265</u>	<u>SAMPLE RATE</u>
0	32	4000 Khz
1	48	4400 Khz
2	64	5000 Khz
3	80	5300 Khz
4	96	5900 Khz
5	112	6500 Khz
6	128	7100 Khz
7	144	8300 Khz
8	160	8900 Khz
9	176	10,500 Khz
10	192	12,500 Khz

Memory location BASE+266. Playback SPEED setting. Same as in recording RATE setting. The same table applies as for RATE except use BASE+266 in the middle column.

Memory locations BASE+267 to BASE+330. This is a list of 64 bytes which gives the order in which the phrases were LEARN(ed). For example, if you LEARN(ed) phrases 3,8,12,45, and 4, in that order, then the memory location starting at BASE+267 will contain 3, followed by 8, then 12, and so on. The memory location, BASE+259, will then contain 5, or the total number of phrases LEARN(ed). The following BASIC program will check memory to see if a phrase has been recorded:

```

10 BASE=PEEK(56)*256
20 INPUT"ENTER PHRASE TO CHECK";P
30 FOR J=0 TO PEEK(BASE+259)-1
40 M=PEEK(BASE+267+J)
50 IF M=P THEN PRINT"FOUND!":GOTO 80
60 NEXT J
70 PRINT"NOT FOUND"
80 END

```

Memory locations BASE+331 to 45567. This is where the actual digitized speech is stored. For the minimum value BASE = 16, we can compute the total available memory for speech as follows: $45567 - 16*256 - 331 = 41140$ bytes.

COMMAND SUMMARY (Recognition)

TRAIN N. Designated word in the range 0 to 31. The number N can be computed. This command constructs a template for the word n using 96 bytes of memory. Re-training the same number creates an average template. Terminate a TRAIN with the RUN/STOP key.

BLANK N. Clears the template for word N for N in the range 0-31. BLANK without a number clears all templates.

RECOG n. Causes the program to wait for an input and attempt recognition by comparing the template made for the input word to those in memory. For n=0 or no number, all 32 templates are scanned. For n=1,2,3, and 4, template numbers 0-7, 8-15, 16-23, and 24-31 are scanned respectively. RECOG 2,4 scans 8-15 and 24-31 only, in that order. The best fit template number is placed in memory location 151. The number 253 is in location 151 if the "Q" key is pushed instead of inputting a word. Numbers 254 and 255 are in location 151 for two kinds of recognition refusals.

MIN n. Sets minimum template difference so that recognition refusal takes place if two templates in memory are too close to the unknown word template to warrant a decision. Range of n is 0-100 with default value of 30. Recognition refusal puts the number 254 in location 151.

MAX n. Sets maximum score allowed so that recognition is refused if no template in memory is a reasonably close match to the unknown word template to warrant a decision. Range is 0-700 with default value of 300. Recognition refusal puts the number 255 in location 151.

TPUT"filename",8. Saves a set of 32 templates in main memory to disk memory.

TFIND"filename",8. Transfers a set of 32 previously saved templates from disk to main memory.

IMPORTANT MEMORY LOCATIONS (Recognition)

Please refer to the section on "THE "VOICE/RECOG" PROGRAM".

SOME GENERAL COMMENTS ON MEMORY USE

Programs which manipulate speech recording, playback, and recognition utilize several addresses in page 0 of

computer main memory. Many of these locations must be retained such that they are not normally available for user written machine language programs. However, contents of these memory locations can be stored and replaced as part of a user program.

Three areas of memory contain various VOICE MASTER programs. A major portion of the machine language program is contained in a 4K memory block between BASIC and kernal ROM's. A general work area for speech recording is contained behind the kernal. Some numerical data for recognition and templates are located behind the BASIC ROM. Stored vocabulary words for playback are located as desired through use of CLEAR n.

It must be realized that the memory map for a computer is quite specific to that computer. Voice Master programs are not transferrable from one computer to another without numerous adjustments being made, even between computers having the same type of microprocessor found in the Commodore 64. Machine language programs, especially for speech recording and playback and disk (or tape) storage and retrieval tasks, make frequent use of utility programs in the kernal.

BLANK

APPENDIX: CASSETTE OPERATING SYSTEM FROM DISK

Because most Commodore 64 (and 128) computers in North America use disk memory, COVOX, Inc. has ceased providing software on cassette. For the benefit of those limited to cassette memory, especially outside North America where cassette systems are more common, we describe here how to construct a cassette operating system from our standard disk. Special software has been incorporated into VOICE MASTER programs to make this possible.

Cassette memory systems may be more sensible to use than disks in many special applications. If reloading software is infrequent, especially in dedicated systems, an idle and inexpensive cassette drive can be justified much more readily than can an equally idle disk drive. Idling power consumption with the cassette will be less as well.

Important programs on the Voice Master disk can be transferred to tape cassette. A special "BACKUP" program (in BASIC) facilitates this for the most important machine language programs, but not for programs in BASIC files. The BACKUP program automatically loads in a machine language program "BU/ML" which handles the transfer. BACKUP is also used to make partial disk copies (machine language programs).

LOAD"BACKUP",8 (return)

which must be RUN after loading. (All BASIC programs must be RUN.) You will get a menu listing of the following machine language programs:

- 1 VM2.1 Record/playback with wedges and FFIND.
- 2 VM2.1R Record, playback, and recognition with
 wedges but no FFIND.
- 3 VM1.0 Recognition with wedges.
- 4 PLAY/FIND Playback (only) without wedges.
- 5 VOICE/RECOG Recognition and template making
 without wedges.
- 6 VM2.1N VM2.1 without FFIND.
- 7 PERFORMANCE.0 Part of music software.
- 8 OBJ.0 Part of music software.
- 9 MEMOVE.0 Part of music software.

These machine language programs remain essentially the same for either disk or tape cassette operation. There may be one difference, however. "VM2.1" contains a special fast loading routine which substantially speeds up loading speech vocabularies from disk. This routine is not applicable for cassette recorders. Nor is it suitable in

countries where power line frequency is other than 60 hertz. Nor is it suitable for other disk drives than the Commodore 1541. Even though included, it will not prevent operation of cassette or other systems because the normal FIND is included as well as the faster version FFIND. The only disadvantage is that "VM2.1" with FFIND capability is about 3 blocks longer than would otherwise be the case.

The program "VM2.1N" is "VM2.1" without the fast find routine. We suggest using this unless FFIND is required (and can properly function).

Programs written in BASIC are loaded from disk in the usual manner and then saved to tape by changing the device number in the SAVE command. However, if a BASIC program loads in some other program from disk, then the number 8 must be embedded in the BASIC program itself--it is no longer just a keyboard operation with LOAD and SAVE. VOICE MASTER demonstration programs use the variable "D" to designate disk or cassette--8 for disk and 1 for cassette. The program PEEK(s) memory location 186 to get the current device number and then this number is used to accomplish the desired device load or save. In the BASIC "BACKUP" program, for example, the machine language program "BU/ML" is loaded from memory through numbered BASIC statements according to the sequence

```
D = PEEK(186)
LOAD"BU/ML",D,1
```

Location (186) contains the number 8 if loading is from a disk (or 9 if from a second disk, etc.) or the number 1 if loading is from a cassette. If "BACKUP" is loaded from disk to main memory, then saved to cassette, and subsequently loaded from cassette to main memory, the original value of 8 for D will automatically become 1!

On a disk, the order in which programs are listed is not important. But unless programs on tape cassette are carefully ordered, the task of loading in programs which call other programs can become very tedious and difficult. For example, the BASIC program "MENU" will load in "VM2.1R", "BAR", and a speech file "GREETING". (Note: All demonstration programs on the VOICE MASTER disk written in BASIC depend on "MENU" having loaded in the same composite "VM2.1R" program.)

YOU CANNOT LIST OR RUN A BASIC PROGRAM THAT USES ONE OF THE VOICE MASTER PROGRAMS UNLESS THE VOICE MASTER PROGRAM RESIDES IN MAIN MEMORY. You can write, save to disk or cassette, or load from disk or cassette without a

VOICE MASTER program in residence--but not LIST or RUN. Otherwise, when the wedged command is confronted, the computer has no means for recognizing it, a syntax error results, and the computer stops in the READY state.

The following catalog lists all of the programs that are on the VOICE MASTER disk. The numbers on the left give the number of blocks of disk memory required for storage.

Main menu and principal machine language programs.

3	"MENU"	Facilitates selection of demonstrations.
14	"VM2.1"	Record/playback with wedges and FFIND.
19	"VM2.1R"	Recognition and Record/playback with wedges. No FFIND.
11	"VR1.0"	Recognition with wedges.
3	"PLAY/FIND"	Record/playback without wedges or FFIND.
8	"VOICE/RECOG"	Recognition without wedges.
11	"VM2.1N"	Same as "VM2.1" but without FFIND.
7	"BACKUP"	For disk and tape backups.
1	"BU/ML"	Loaded from "BACKUP".

Recording and playback demonstration programs.

8	"DEMO"	Loads in "BAR" and "GREETING".
3	"BAR"	For calibration and display.
16	"GREETING"	Speech vocabulary for "DEMO".
14	"CLOCK".	
7	"CALCULATOR".	

Word recognition demonstration programs.

4	"RECOG-DEMO".	
30	"BLACKJACK".	
41	"VOICE"	The dealer's voice in "BLACKJACK".
6	"VOICE-CREATOR"	Users voice for dealer.
9	"CALCULATOR-2".	

Music programs.

27	"HUM-ALONG"	For computer output with voice or whistle input. Independent of speech.
8	"PERFORMANCE.0"	Loaded from "HUM-ALONG".
5	"PRESET-1"	Special instructions as part of "BACKUP" show how to load this file.
11	"COMPOSER"	Music. Independent of speech. It loads the following programs.
26	"OBJ.0"	Loaded from "COMPOSER".
9	"NOTE"	Loaded from "COMPOSER".
21	"PRINT-SONG"	Not needed if printout from composer is not desired. Not applicable unless COMPOSER also present.
1	"MEMOVE.0"	Loaded from "PRINT-SONG".

Speech and template files in the foregoing list are loaded and saved using the VOICE MASTER commands FIND, TFIND, PUT, and TPUT. Transfer from disk to tape requires changing the number "8" to "1" for SAVE as a keyboard command. The same procedure applies for any program written in BASIC.

Any of the listed programs which show "VM2.1" as a requirement can work equally well (less fast find) with "VM2.1N" or "VM2.1R". Recognition programs will not operate with "VM2.1" or "VM2.1N". If the BASIC program is loaded and RUN from tape cassette, and if the necessary VOICE MASTER program has not yet been loaded, the program stops almost immediately and informs which assembly language program should next be loaded. When the tape is started, it will keep going until the properly named program is found. Then you must go back and load in the original BASIC program. Proper ordering of programs on the tape (or use of separate short tapes) can save time.

Note: When loading "VM2.1R" from tape cassette, you may get "?LOAD ERROR". This is normal. Ignore it.

Music program: "HUM-ALONG". This program is in BASIC. Load and save in the normal way. Do not rewind tape after saving. Then load VOICE MASTER "BACKUP" and follow instructions to save "PERFORMANCE.0". Do not rewind tape. Then re-load "HUM-ALONG" from disk and RUN it. It will ask for "PRESET-1" and you can then load this from disk and save on tape.

Music program: "COMPOSER". This program is in BASIC. Load and save in the normal way. Do not rewind tape after saving. Then load VOICE MASTER "BACKUP" and follow instructions to save "OBJ.9". Do not rewind tape. Then use "BACKUP" a second time to save "NOTE" on cassette.

Music program: "PRINT-SONG". This BASIC program will print out a song that has been written with "COMPOSER" provided that the song has been saved or remains in main memory. "PRINT-SONG" is saved on tape as for any BASIC program. Do not rewind tape. Then use "BACKUP" to save "MEMOVE.0" on tape. When "PRINT-SONG" is subsequently loaded from tape and RUN, it will start the tape and load "MEMOVE.0".

BLANK



COVOX inc.

675-D Conger Street
Eugene, OR 97402

BULK RATE
U.S. Postage
PAID
Permit No. 38
Eugene, OR

RETURNED POSTAGE GUARANTEED