D1 D2	1N914 diodes	2
D3	Red LED	1
	9V Battery Holder	1
P1	PC mount microphone jack	1
P5	2-position header	1
	Jumper	1
	Headset Microphone	1
	3V Coin Battery	1
2/56 Hex nuts, Screws & Lock washers		2 each
7-pin headers (male and female)		1 each
10-pin headers (male and female)		1 each

SR-06/SR-07
Speech Recognition Kit
Construction Manual
& User Guide
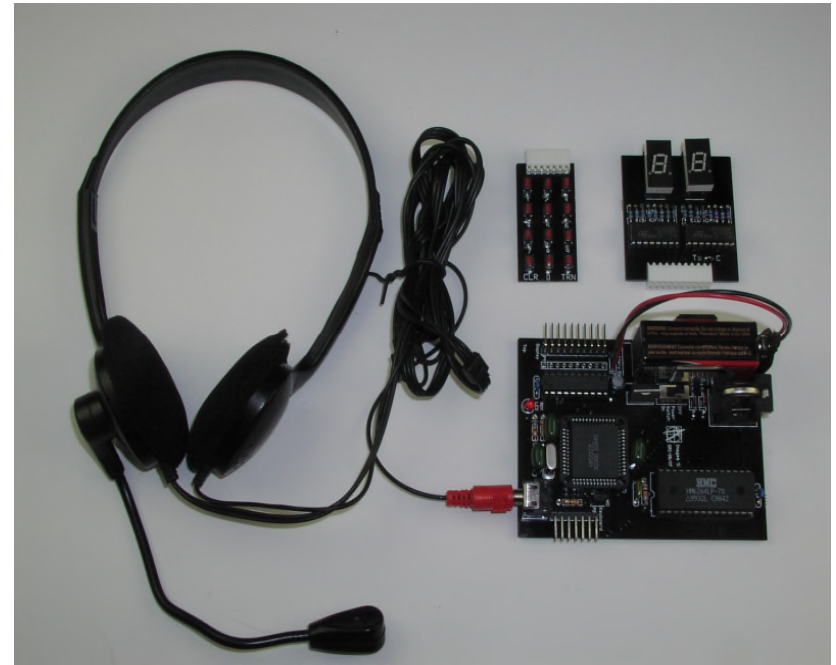
Images SI Inc.
Staten Island NY 10312
718.966.3694 voice
718.966.3695 fax
http://www.imagesco.com

Table of Contents

Maximum word length 1.92 seconds (20 word)
Microphone support
Manual and CPU modes available
Response time less than 300 milliseconds
5V power supply

More information on the HM2007 chip is available in the HM2007 data booklet (DS-HM2007).

Parts List:

Placement	Item	Quantity
	PCB	3 pieces
Keypad	Push-button Switches	12
U4	HM2007 PLCC	1
	52-pin socket	1
U1	7805 Voltage Regulator	1
U2	74LS373	1
	20-pin socket	1
U3	SRAM 8K X 8	1
	28-pin socket	1
U4 U5	4511	2
	16-pin socket	2
U6 U7	220ohm 1/8W Resistors	16
U8 U9	7-Segment Displays	2
	14-pin socket	2
X1	XTAL 3.57 MHz	1
S1	Toggle Switch	1
BT1	9V Battery Snap	1
BT2	Coin Batter Holder	1
R1	100K 1/4W Resistor	1
R2	6.8K 1/4W Resistor	1
R3	22K 1/4W Resistor	1
R4	330ohm 1/4W Resistor	1
C1	100 uF Capacitor	1
C2 C3	22 pF Capacitors	2
C4	.0047 uF Capacitor	1
C5 C6 C7	.01 uF Capacitors	3

require speaker independent voice recognition systems.

Recognition Style

In addition to the speaker dependent/independent classification, speech recognition also contends with the style of speech it can recognize. They are three styles of speech: isolated, connected and continuous.

Isolated. Words are spoken separately or isolated. This is the most common speech recognition system available today. The user must pause between each word or command spoken.

Connected. This is a half way point between isolated word and continuous speech recognition. It permits users to speak multiple words. The HM2007 can be set up to identify words or phrases 1.92 seconds in length. This reduces the word recognition dictionary number to 20.

Continuous. This is the natural conversational speech we use to in everyday life. It is extremely difficult for a recognizer to sift through the sound as the words tend to merge together. For instance, "Hi, how are you doing?" to a computer sounds like "Hi,.howyadoin" Continuous speech recognition systems are on the market and are under continual development.

More On The HM2007 Chip

The HM2007 is a CMOS voice recognition LSI (Large Scale Integration) circuit. The chip contains an analog front end, voice analysis, regulation, and system control functions. The chip may be used in a stand alone or CPU connected.

Features:
      Single chip voice recognition CMOS LSI
      Speaker dependent
      External RAM support
      Maximum 40 word recognition (.96 second)

Introduction:

The speech recognition kit is a complete easy to build programmable speech recognition circuit. Programmable, in the sense that you train the words (or vocal utterances) you want the circuit to recognize. This kit allows you to experiment with many facets of speech recognition technology.

Features of the kit include:

Self-contained stand alone speech recognition circuit
User programmable
40 or 20 word vocabulary
Multi-lingual
Non-volatile memory back up
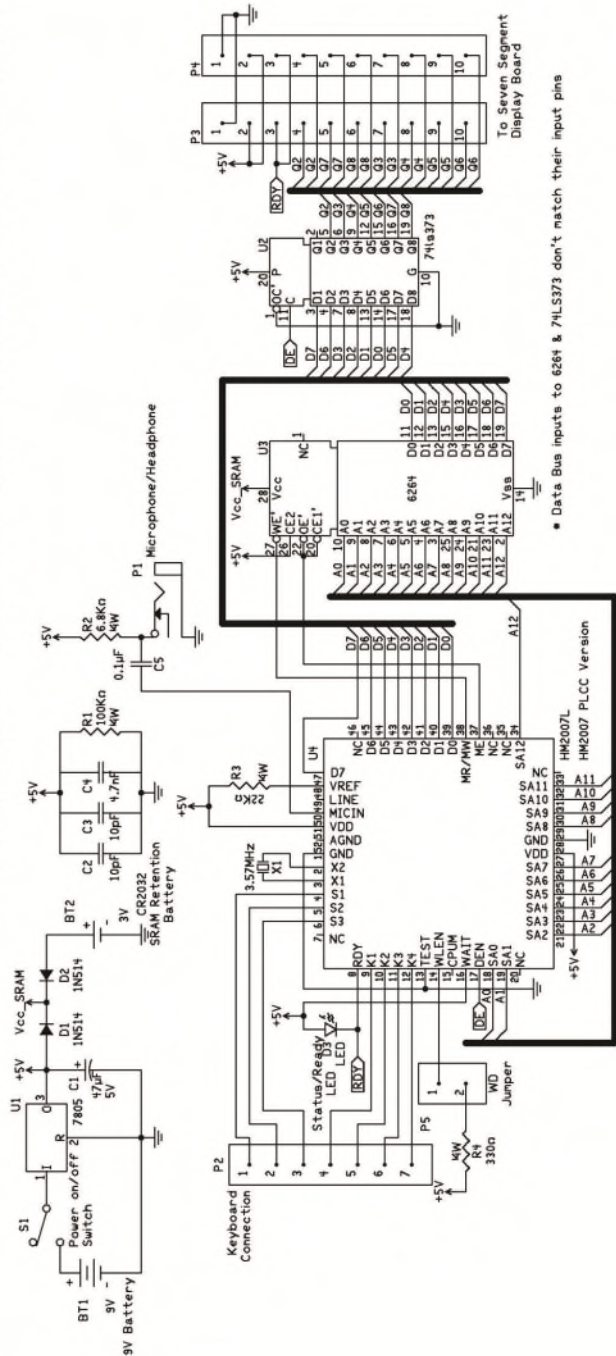Easily interfaced to control external circuits & appliances

Speech recognition will become the method of choice for controlling appliances, toys, tools and computers. At its most basic level, speech controlled appliances and tools allow the user to perform parallel tasks (i.e. hands and eyes are busy elsewhere) while working with the tool or appliance.

The heart of the circuit is the HM2007 speech recognition IC. The IC can recognize either 40 words, each word a length of .96 seconds or 20 words, each word a length of 1.92 seconds.

Applications

There are several areas for application of voice recognition technology.

      Speech controlled appliances and toys
      Speech assisted computer games
      Speech assisted virtual reality
      Telephone assistance systems
      Voice recognition security
      Speech to speech translation

unstructured and efficient methods. The way in which people typically communicate verbally.

Learning To Listen

The ability to listen to one person speak among several at a party is beyond the capabilities of today's speech recognition systems. Speech recognition systems can not (as of yet) separate and filter out what should be considered extraneous noise.
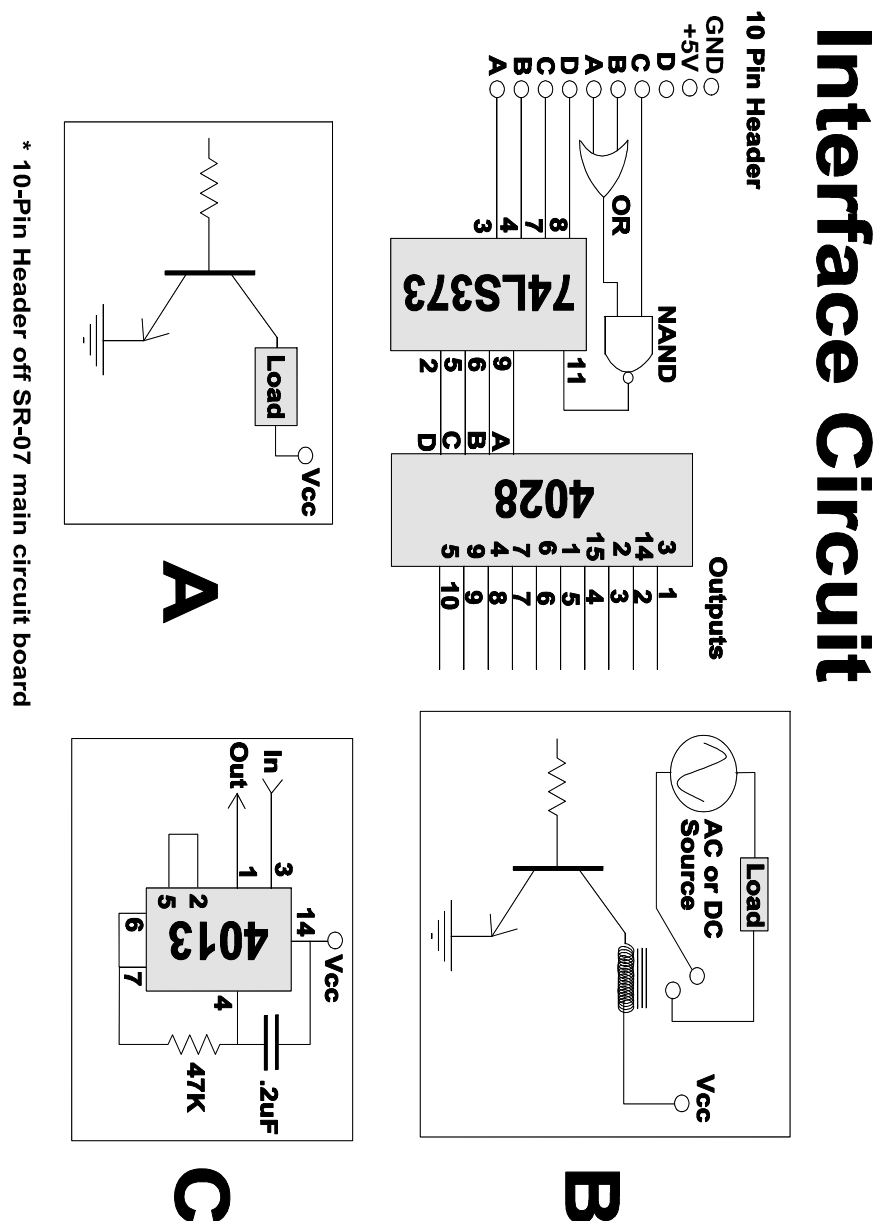
Speech recognition is not understanding speech. Understanding the meaning of words is a higher intellectual function. Because a circuit can respond to a vocal command doesn't mean it understands the command spoken. In the future, voice recognition systems may have the ability to distinguish nuances of speech and meanings of words, to "Do what I mean, not what I say!"

Speaker Dependent / Speaker Independent

Speech recognition is divided into two broad processing categories; speaker dependent and speaker independent.

Speaker dependent systems are trained by the individual who will be using the system. These systems are capable of achieving a high command count and better than 95% accuracy for word recognition. The drawback to this approach is that the system only responds accurately only to the individual who trained the system. This is the most common approach employed in software for personal computers.

Speaker independent is a system trained to respond to a word regardless of who speaks. Therefore the system must respond to a large variety of speech patterns, inflections and enunciation's of the target word. The command word count is usually lower than the speaker dependent however high accuracy can still be maintain within processing limits. Industrial applications more often

The schematic for the SR-07 Speech Recognition Circuit is shown in figure 1. The SR-07 utilizes three separate printed circuit boards (pcb). The components are mounted on the top side of each pc board. The top sides of the boards have white silk screen component drawings. The components are soldered on the opposite side of the pc board. After soldering the component to the board any excess wire is clipped off.

Chip Installation

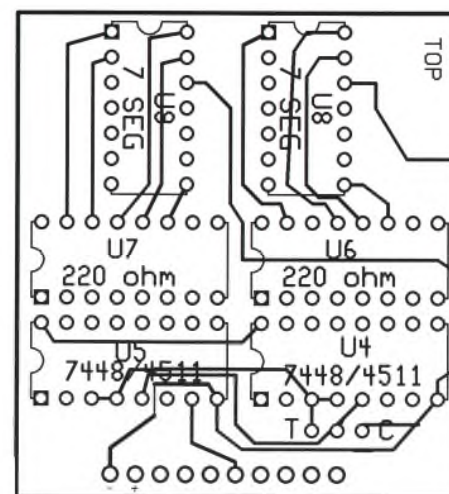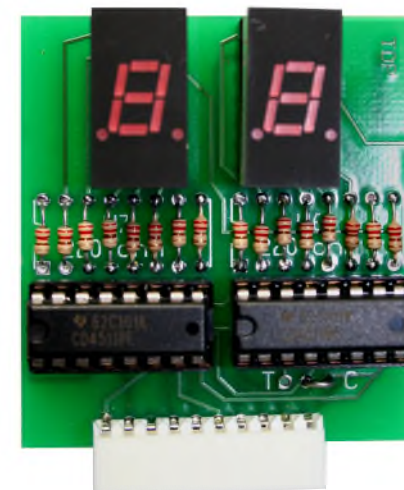When installing integrated circuit (IC) chips, begin by first identifying the top of the chip. The top of the chip has a marker, many times it is a half circle cutout. Sometimes it is a small mark identifying pin 1 on the IC. In both cases the marks show us the top of the IC chip. Orientated the top of IC chips with the white silk screen drawings of the components on the top of the pc board (usually a half circle cutout) or on the parts placement drawings and install the IC into their socket.
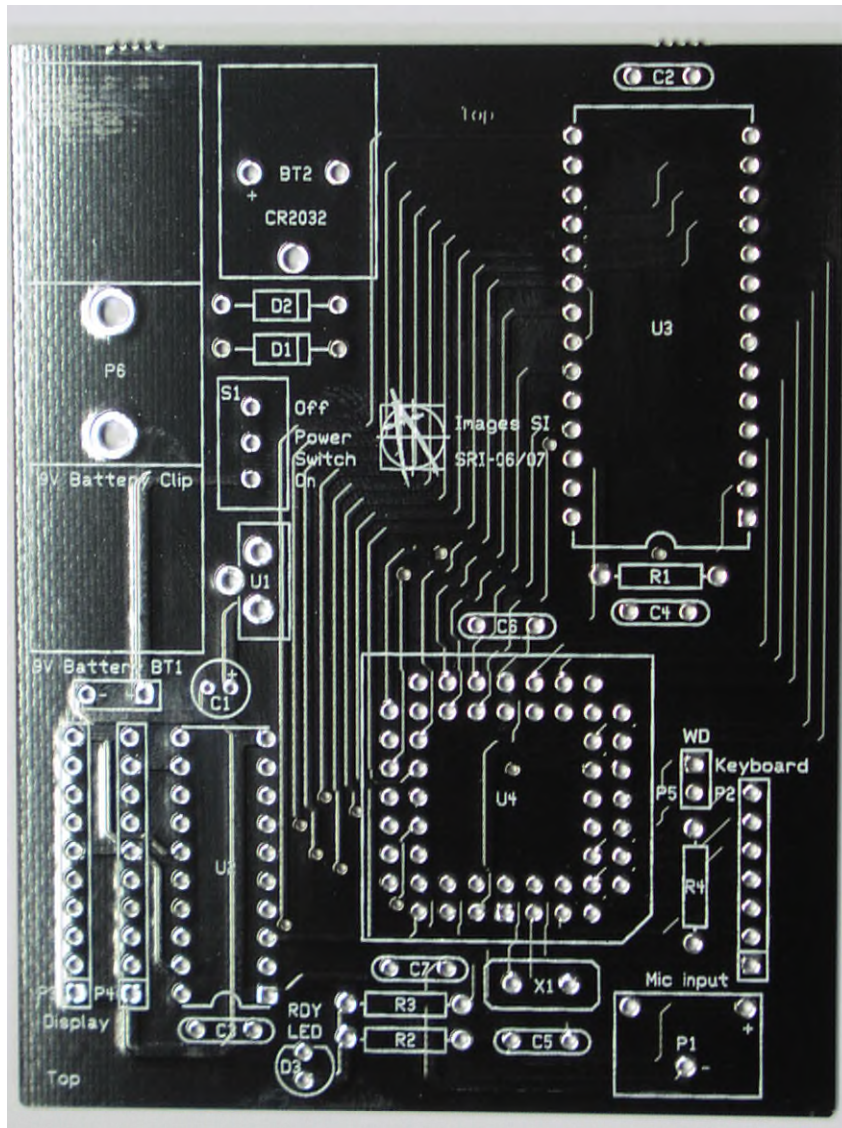
Display Board:

We start construction with the display board, see figure 2.

Figure 2

Finished Display Board

Figure 5

Interface Circuit

* 10-Pin Header off SR-07 main circuit board

A

B

C

Figure 3

OK, lets see how this works in the real world. Lets assume you have the power to a printer connected through the speech recognition circuit and a 4013 controlled switch or relay. The target word is "printer". When you want to turn the printer on you use the command word "printer". The circuit recognizes the word and applies power to the printer. At this point you can also turn on (or off) any other circuit connected to the speech board, because when the high signal goes low the 4013 keeps its output high. When you are ready to turn off the power to the printer all you have to do is repeat the command "printer". The second time the line is brought high, the output of the 4013 is brought low.

The same command is used to turn the unit on and off. Any of the other lines can be turned on or off without affecting the status of the other output lines.

Voice Security System

This circuit isn't designed for a voice security system in a commercial application, but that should not prevent anyone from experimenting with it for that purpose. A common approach is to use three or four keywords that must be spoken and recognized in sequence in order to open a lock or allow entry.

CPU Mode

The HM2007 speech recognition chip has a CPU mode to be used when connected to a host computer system or microcontroller. To interface the HM2007 to a host computer requires the writing of driver software as well as designing and building the hardware interface to the computer data buss.

Aural Interfaces

It's been found that mixing visual and aural information is not effective. Products that require visual confirmation of an aural command grossly reduces efficiency. To create an effective AUI products need to understand (recognize) commands given in an

Error Codes

The decoding circuit must recognize the word numbers from error codes. So the circuit must be designed to recognize error codes 55, 66 and 77 and not confuse them with word spaces 5, 6 and 7. This is accomplished using an OR gate and a NAND gate connected to the MSB. Whenever an error code is generated, the word number is ignored.

Interface Circuit

Figure 5 is the schematic for the interface circuit. The circuit connects to the 10 pin Right Angle interface header on the circuit board. This header is also used for the Digital Display board

The 4028 has ten output lines. Whatever number is displayed on the LSD the corresponding line number off the 4028 will be brought high. The high signal from the 4028 can be connected to a NPN transistor to control a DC load as shown in box A, or control an AC or DC load using a simple relay as shown in box B.

The disadvantage to using a simple set up like this is that only one switch out of 10 may be turned on at any given time. This doesn't make for a very good system, so a solution must be found that allows one to turn on or off any line without changing the status of any other line.

This can be implemented by inserting a flip-flop, shown in box C, between the 4028 and the NPN transistor. The 4013 IC contains two flip-flops, only one is shown in the box C drawing. The flip-flop acts like simple memory. When the input line is brought high, its output line goes high, turning on the NPN transistor. When the output line is brought low, the output line still stays high. When the flip-flop receives a second high signal on its output line it brings the output low.

Mount and solder 16 (220 ohm) resistors, color bands red-red-brown gold or silver. Next solder two 14-pin sockets for the LED display ICs (U8 and U9). Install the LED displays into the sockets, the dots on the display chips face the bottom of the pc board. Next mount and solder the two 16-pin sockets for the 4511 IC's (U4 and U5) making sure to orientate the sockets properly. Install the 4511 ICs in their appropriate IC sockets making sure to orientate the ICs properly. Just below U4 there are three solder pads in a row. Solder a jumper wire from the center pad to the right pad marked with a "C". Finish the display board by mounting and soldering the 10-pin female header to the PC board.

Main Circuit Board:

The pc board layout of the main board is shown in figure 3. Begin construction of this board by mounting and soldering the three IC sockets. The HM2007 PLCC uses a 52 pin square socket identified on the pcb as U1. The 8K static RAM uses a 28 pin socket identified as U2. The 74LS373 uses a 20 pin socket identified as U3. Next mount and solder resistor R1, it has a nominal value of 100K. Its color bands brown, black, yellow, gold. Solder resistor R2, it has a nominal value of 6.8K. Its color bands are blue, grey, red and gold. Then solder resistor R3, it has a nominal value of 22K (color bands red, red, orange and gold). Next solder resistor R4, it has a nominal value of 330 ohms (color bands orange, orange, brown, gold).

Mount and solder diodes D1 and D2. Make sure the black band faces the correct direction as shown on the drawings. Next mount and solder the 3.57 MHz crystal. It is identified as XTAL on the parts placement drawing. Mount and solder the red LED next. The short lead of the LED should be aligned with the flat side of the silkscreen circle marked LED. Mount & solder capacitors C1 to C7. Capacitors C2 and C3 are small 22 pf capacitors. C5, C6 & C7 are .1 uF capacitors, C1 is 47 to 100 uF capacitor and C4 is .0047 uF capacitor. Please note C1 is identified as a 47 uF capacitor, but any value between 47 and

SR-07 Main Circuit Board
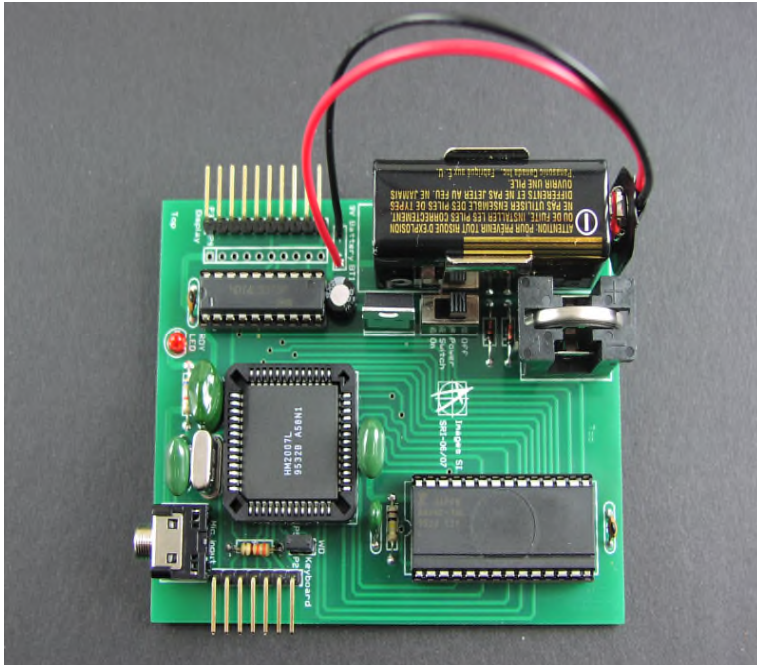
100 uF may be substituted in the kit.
Mount & solder the 7805 voltage regulator and on-off slide switch. Next mount and solder the microphone jack and button battery holder and 9-volt battery cap. Keep the wires on the 9-volt battery cap short, about 1.5" long.
Mount and solder the 10-pin right angle header in the upper left hand corner of the board identified as R1. Mount and solder the 7-pin right angle header in the lower left corner of the board, Mount and solder a 2-pin header in the WD location next to R4. Install the integrated circuits in their appropriate IC sockets making sure to orientate the IC's properly.

Keypad:

The keypad is constructed using 12 normally open momentary contact switches. Place each switch in its mounting position and bend the leads inward to secure the switch to the PCB for soldering. After mounting and soldering the 12 keypad switches to the

Rhyming words

Rhyming words are words that sound alike. For instance the words cat, bat, sat and fat sound alike. Because of their like sounding nature they can confuse the speech recognition circuit. When choosing target words for your system do not use rhyming words.

The Voice With Stress & Excitement

Stress and excitement alters ones voice. This affects the accuracy of the circuit's recognition. For instance assume you are sitting at your workbench and you program the target words like fire, left, right, forward, etc., into the circuit. Then you use the circuit to control a flight simulator game, Doom or Duke Nukem. Well, when you're playing the game you'll likely be yelling "FIRE!…Fire!...FIRE!!...LEFT…go RIGHT!". In the heat of the action you're voice will sound much different than when you were sitting down relaxed and programming the circuit. To achieve a higher accuracy word recognition one needs to mimic the excitement in ones voice when programming the circuit.

These factors should be kept in mind to achieve the high accuracy possible from the circuit. This becomes increasingly important when the speech recognition circuit is taken out of the lab and put to work in the outside world.

Interfacing The Circuit To The Outside World

The circuit design for interfacing the speech recognition system to the outside world controls ten switches. This design idea fits with the robust speech recognition system discussed previously. While the effective vocabulary drops from forty to ten words, we gain in a more robust and accurate system.

Simulated Independent Recognition

The speech recognition system is speaker dependant, meaning that the voice that trained the system has the highest recognition accuracy. But you can simulate independent speech recognition.

To make the recognition system simulate speaker independence one uses more than one word space for each target word. Set the SR-07 for a 40 word vocabulary. Now we use four word spaces per target word. Therefore we obtain four different enunciation's of each target word. (speaker independent)

The four word spaces are chosen to minimize software and hardware interfaces into the circuit. To accomplish this the four word spaces are chosen so that they all have the same Least Significant Digit (LSD). Doing this the words can be recognized by just decoding the least significant digit (number) on the digital display.

Using this procedure the word spaces 01, 11, 21 and 31 are allocated to the first target word. The Most Significant Digit (MSD) is dropped by the interfacing circuits. By only decoding only the LSD number, in this case 1 of "X1" (where X is any number) we can recognize the target word.

We continue do this for the remaining word space. For instance, the second target word will use the word spaces 02, 12, 22 and 32. We continue in this manner until all the words are programmed.

If you are experimenting with speaker independence use different people when training a target word. This will enable the system to recognize different voices, inflections and enunciation's of the target word. The more system resources that are allocated for independent recognition the more robust the circuit will become.

If you are experimenting with designing the most robust and accurate system possible, train target words using one voice with different inflections and enunciation's of the target word.

Figure 4
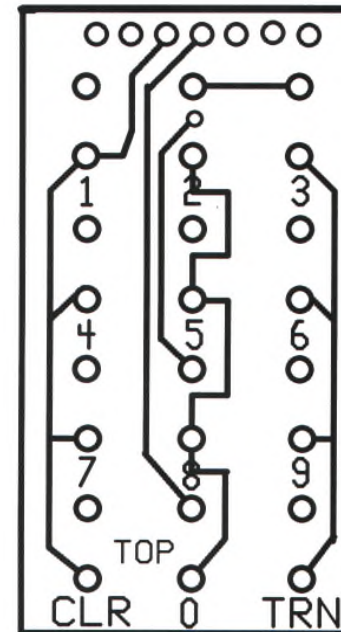
top of the keypad PCB, connect the 7-pin female header to the bottom of the keypad pcb.

Non-Volatile Memory Back-up

The PC mounted coin battery holder holds a 2032 coin battery, which supplies backup for the SRAM. This allows the word patterns to be retained in memory when the main circuit is turned off.

Selecting Vocabulary Size and Word Length

The default vocabulary and word configuration for the circuit is 40 words, each with a length of .96 seconds. If you wish to change this to 20 word with a length of 1.92 seconds place a jumper on the two-pin WD header.

If you do not need the 40 word vocabulary, it is suggested you configure the circuit for the 20 word vocabulary as this configuration usually provides a better recognition accuracy.

Keypad

Using The Speech Recognition Circuit

The keypad and digital display are used to communicate with and program the HM2007 chip. Plug the digital display into the 10-pin header on the main circuit board. Plug the keypad into the 7-pin
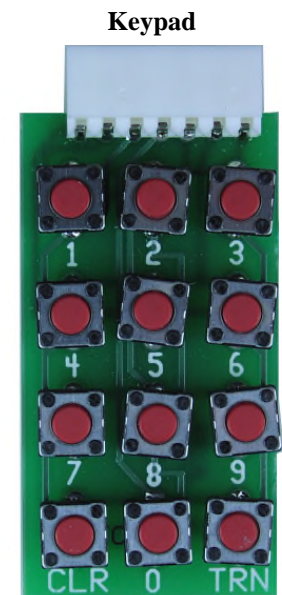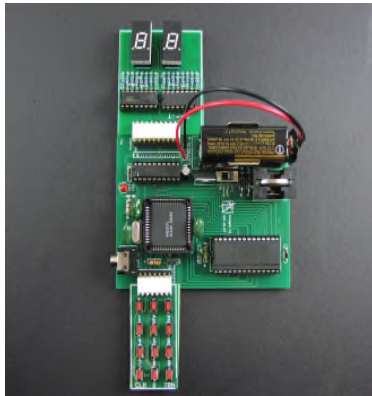
header on the main circuit board. Plug the headset microphone into the microphone jack. Adjust the microphone so that it is po- sition about 1" away from your mouth.

Keypad Use:

The keypad is made up of 12 normally open momentary contact switches.

Finished SR-07 Circuit


```
1    2    3

4    5    6

7    8    9

CLR  0  TRN
Clear     Train
```

The CLR key equals Clear and the TRN key equal Train.

When the circuit is turned on, "00" is on the digital display, the red LED (READY) is lit and the circuit waits for a command.

Training Words for Recognition

To Train:

Press "1" (display will show "01" and the LED will turn off) on the keypad, then press the TRN key (the LED will turn on) to place circuit in training mode, for word one.

Say the target word into the headset microphone clearly. The circuit signals acceptance of the voice input by blinking the LED off then on. The word (or utterance) is now identified as the "01" word. If the LED did not flash, start over by pressing "1" and then "TRN" key.

You may continue training new words in the circuit. Press "2" then TRN to train the second word and so on. The circuit will accept and recognize up to 40 words (numbers 1 through 40). It is not necessary to train all word spaces. If you only require 10 target words that's all you need to train.

Testing Recognition:

Repeat a trained word into the microphone. The number of the word should be displayed on the digital display. For instance, if the word "directory" was trained as word number 25, saying the word "directory" into the microphone will cause the number 25 to be displayed.

Error Codes:

The chip provides the following error codes.

> 55 = word to long
> 66 = word to short
> 77 = no match

Clearing Memory

To erase all words in memory press "99" and then "CLR". The numbers displayed are "19", this is not an error. The numbers will quickly scroll by on the digital display as the memory is erased.

Changing & Erasing Words

Trained words can easily be changed by overwriting the original word. For instances suppose word six was the word "Capital" and you want to change it to the word "State". Simply retrain the word space by pressing "6" then the TRN key and saying the word "State" into the microphone.
If one wishes to erase the word without replacing it with another word press the word number (in this case six) then press the CLR key. Word six is now erased.