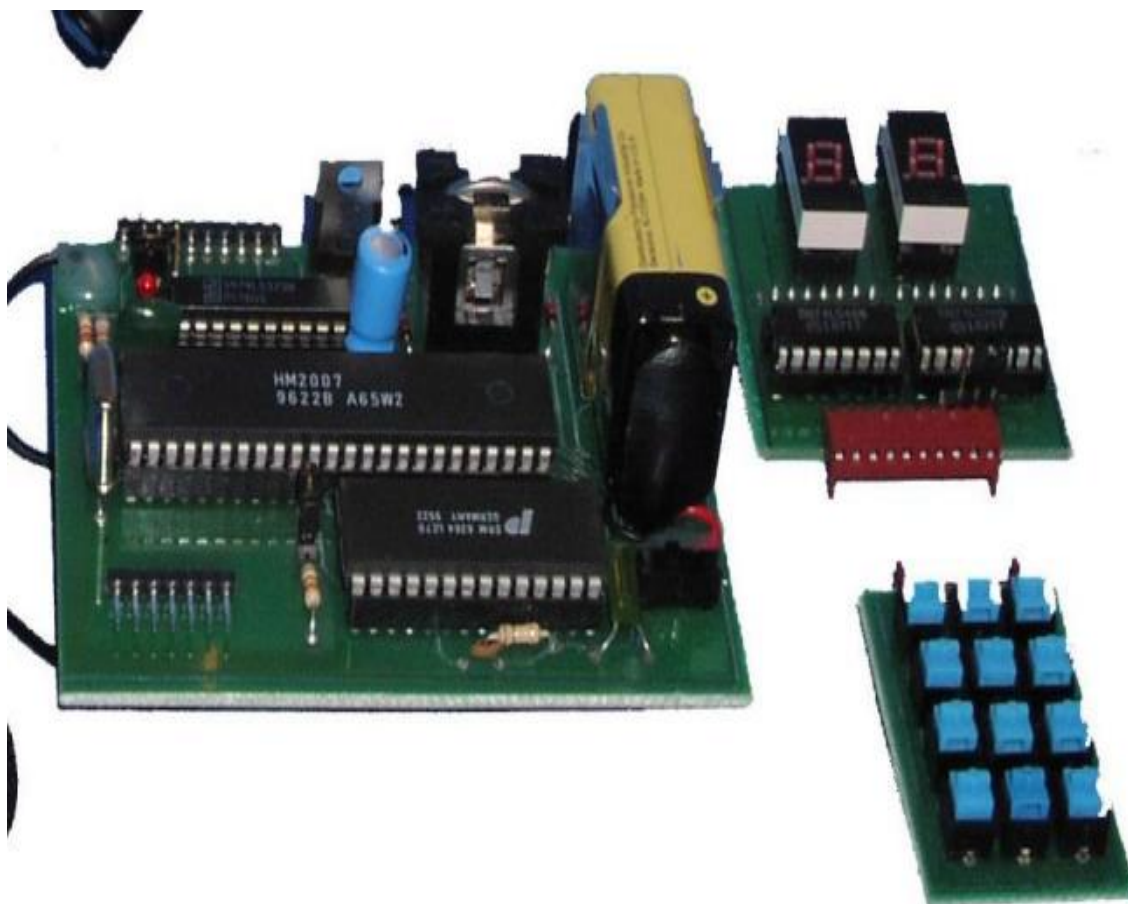




See this project on [Talking Electronics Interactive](http://TalkingElectronicsInteractive.com) website

## Speech Recognition

In this article we build a speech recognition project. The two articles to follow will be speech recognition interfaces. The first interface will connect to a robotic arm (OWI's 007 Robotic Arm Trainer) to create a speech controlled robotic arm. The second interface will be a general purpose PC board that will be easily interfaced to your own projects.



THE THREE SPEECH RECOGNITION MODULES

This speech project design takes a modular approach. Both the keypad and digital display are removable from the main circuit board. Once the circuit is trained and tested, the keypad and display are not needed. They only need to be reconnected for changing words and retraining. Removing the keypad and display from the main circuit board simplifies embedding the speech recognition board into your design. The two speech interfaces we will build later, both plug into the digital display connector on the main circuit board.

## **Why Speech Recognition?**

In the near future, speech recognition will become the method of choice for controlling appliances, toys, tools, computers and robotics. There is a huge commercial market just waiting for this technology to mature.

To control and command an appliance (computer, VCR, TV security system, etc.) by speaking to it, will make it easier to use, while increasing the efficiency and effectiveness of working with that device. At the most basic level, speech recognition allows the user to perform parallel tasks, (i.e. hands and eyes are busy elsewhere) while continuing to work with the computer or appliance.

Our circuit is a stand-alone trainable speech recognition circuit that may be interfaced to control just about anything electrical, such as; appliances, robots, test instruments, VCR's TV's, etc. The circuit is trained (programmed) to recognize the words you want it to recognize. The unit can be trained in any language and even non-languages such as grunts, birdcalls and whistles. The entire speech recognition circuit is available as a kit (SR-07) or may be hardwired together in accordance with the schematic.

This circuit allows you to experiment with many facets speech recognition technology. The heart of the circuit is the HM2007 speech recognition integrated circuit. This chip provides the options of recognizing either forty .96 second words or twenty 1.92 second words. A jumper on the main circuit board selects either the .96 second word length (40 word vocabulary) or the 1.92 second word length (20 word vocabulary). I typically use the 1.92 second option because I found the recognition to be more accurate.

The HM2007 stores the "trained" word patterns used for recognition in external memory. For memory, the circuit uses an on board 8K X 8 static RAM. The main board has a coin battery holder that provides backup power to the static ram when the main circuit is turned off. This keeps all the trained words safely stored in memory (sram) so the circuit does not have to be retrained every time it is turned on. A fresh coin battery provides years of memory protection.

## **Applications:**

- Command and control of appliances and equipment:
- Telephone assistance systems,
- Data entry
- Speech controlled toys,
- Speech and voice recognition security systems.

## **Software Approach**

Currently, most speech recognition systems available today are programs that use personal computers equipped with a sound card. These memory-resident programs operate continuously in the background of the computers operating system (Windows, OS/2, etc.), allowing the speech recognition program to be used with other programs like Word or Excel. There is a noticeable slow-down in the operation and function of the computer when the memory resident voice recognition program is enabled. The memory resident programs add to the processing overhead of the computer's CPU.

From a commercial aspect, the disadvantage in this approach is the necessity of a computer. While these speech programs are impressive, it is not economically viable for manufacturers to add full blown computer systems to control a washing machine or VCR.

## Learning to Listen

We take our ability to listen for granted. For instance, we are capable of listening to one person speak among several at a party. We sub-consciously filter out the extemporaneous conversations and sound. This filtering ability is beyond the capabilities of today's speech recognition systems.

Speech recognition is not speech understanding. Understanding the meaning of words is a higher intellectual function. Because a computer can respond to a vocal command does not mean it understands the spoken command. Voice recognition system will one day have the ability to distinguish linguistic nuances and meaning of words, to "Do what I mean, not what I say!"

## Speaker Dependent / Speaker Independent

Speech recognition is classified into two 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 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 maintained within processing limits. Industrial requirements more often need speaker independent voice systems, such as the AT&T system used in the telephone systems.

## Recognition Style

Speech recognition systems have another constraint concerning the style of speech they can recognize. There are three styles of speech: isolated, connected and continuous.

**Isolated** speech recognition systems can just handle words that are spoken separately. This is the most common speech recognition systems available today. The user must pause between each word or command spoken. The speech recognition circuit is set up to identify isolated words of .96 second lengths.

**Connected** is a half way point between isolated word and continuous speech recognition. It allows 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 vocabulary to 20.

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

## Speech Recognition Circuit

The speech recognition circuit (SR-07) uses a simple keypad and digital display to communicate with and program the HM2007 chip.

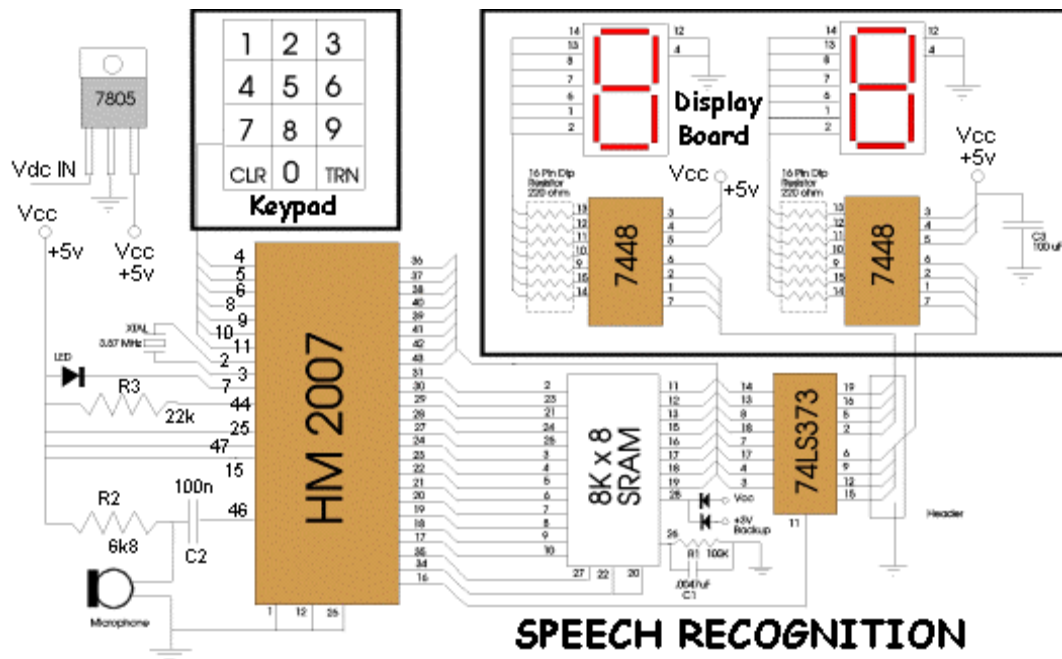


Figure 1

## Keypad

The keypad is made up of 12 switches.

1 2 3

4 5 6

7 8 9

\* 0 #

Clear Train

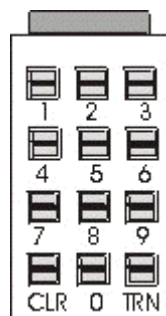
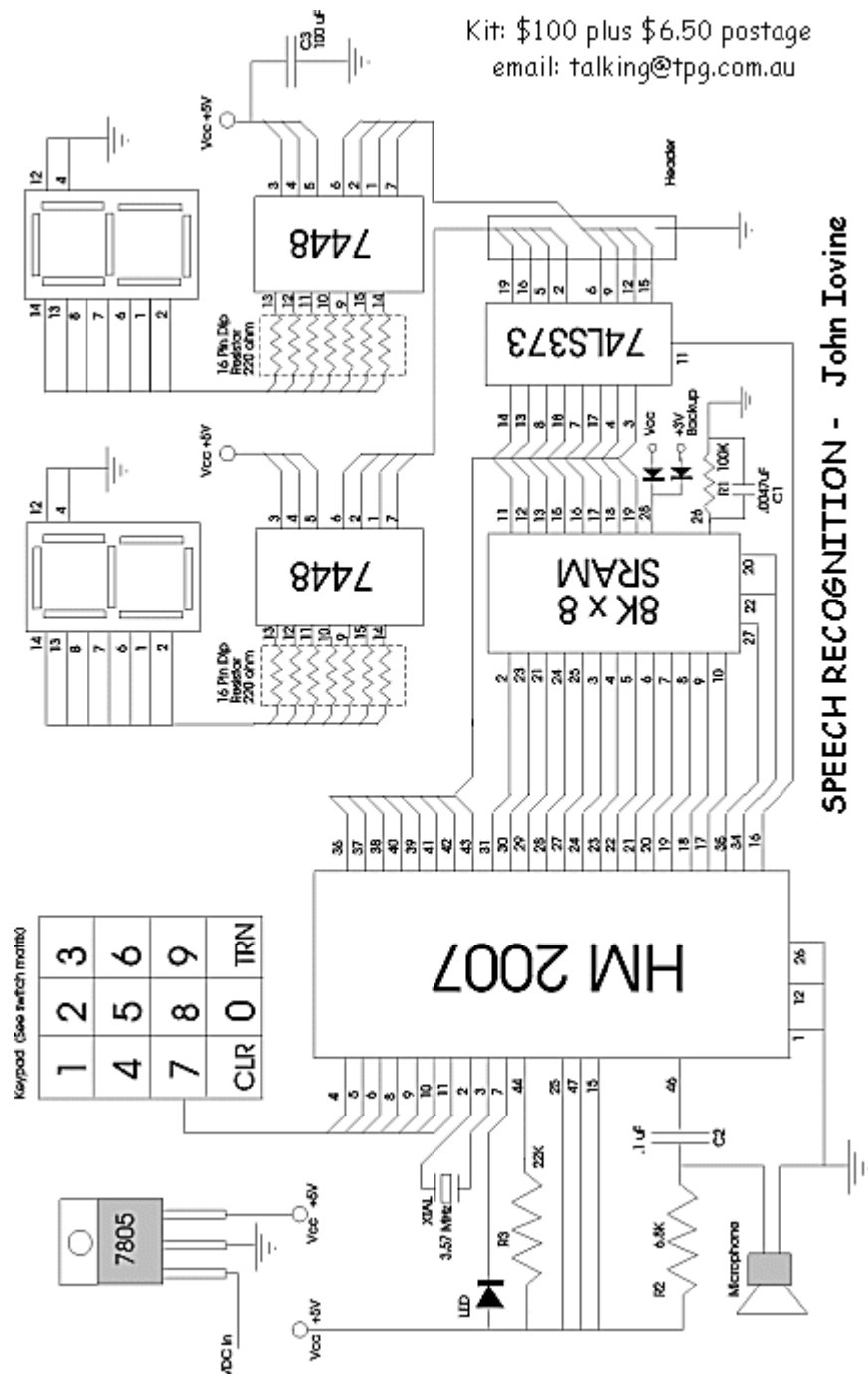


Figure 2



## SPEECH RECOGNITION - John Iovine

When the circuit is turned on, the HM2007 checks memory (static RAM) status. If successful, the board displays "00" on the digital display and lights the red LED (READY). In the "Ready" state, the circuit is listening for a verbal word to recognize or may be programmed (trained).

## Programming the HM2007

### To Train

To train the circuit, begin by pressing the word number you want to train on the keypad. The circuit can be trained to recognize either 40 (one-second) words or 20 (two-second) words. This option is selectable by setting a jumper on the main circuit board. Use any numbers between 1 and the range you've chosen (20 or 40). For example press the number "1" to train word number 1. When you press the number(s) on the keypad the red LED will turn off. The number is displayed on the digital display. Next press the "#" key for train. When the "#" key is pressed it signals the chip to listen for a training word and the red LED turns back on. Now speak the word you want the circuit to recognize into the microphone. The LED should blink off momentarily, this is a signal that the word has been accepted.



**Figure 3: SR-07 Kit ready to program (head set not shown)**

Continue training new words using the procedure outlined above. Press the "2" key then "#" key to train the second word and so on. You do not have to enter the maximum amount of words into memory to use the circuit. You can train as many words as you require.

### Testing Recognition

The circuit is continually listening. Repeat a training 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 5. Saying the word "directory" into the microphone will cause the number 5 to be displayed on the digital display.

### Error Codes:

The chip provides the following error codes.

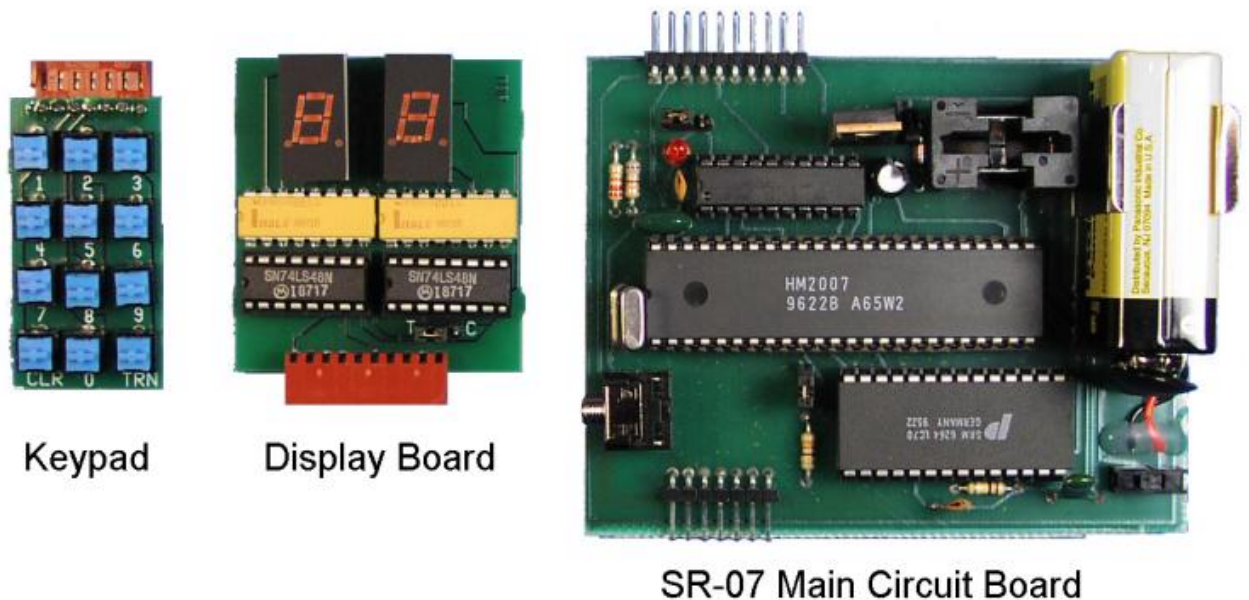
- 55 = word too long
- 66 = word too short
- 77 = word no match

## Clearing the memory

To erase all the words in the RAM memory (Training) press "99" on the keypad then press the "x" key. The display will scroll through the numbers 1-40 (or 1-20) quickly, clearing out the memory. To erase a single word space press the number of the word you want to clear, then press the "x" key.

## Circuit Construction

The schematic for the circuit is shown in figure 1. The kit includes all the components, headset microphone plus the three PC boards. The construction details are provided in the construction booklet that's included with the kit, and I will not repeat them here. It generally amounts to mounting and soldering the components on to the three PC boards, see figure 3.



**Figure 4 SR-07 consists of three modular boards.**

## Independent Recognition System

This speech recognition circuit allows you to experiment with and mimic speaker independent systems. Even though, the speech recognition system is sold as speaker dependent.

To train the system for speaker independent recognition (Multi-user) use the following technique. We will use four word spaces for each target word. Let's arrange the words so that the words can be recognized by just decoding the least significant digit (number) on the digital display.

To accomplish this word, spaces 01, 11, 21 and 31 are allocated to the first target word. By only decoding the least significant digit number, in this case 1 of "X" "1" (where X is any number 0 - 3) we can recognize the target word.

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

If possible, use a different person speaking the word. This will enable the system to recognize different voices, inflections and enunciations of the target word. The more system resources that are allocated for independent recognition the more robust the circuit will become.

There are certain caveats to be aware of. First you are trading-off word vocabulary number for speaker independence. The effective vocabulary drops from forty words to ten words.

The decoding circuit that recognizes the word number and performs a function must be designed to recognize error codes 55, 66 and 77 and not confuse them with word spaces 5, 6 and 7. Our interface circuit does this.

### **Voice Security System**

This HM2007 wasn't designed for use in a voice security system. But this doesn't prevent you from experimenting with it for that purpose. You may want to use three or four keywords that must be spoken and recognized in sequence in order to activate a circuit that opens a lock or allows entry.

### **Coming Next**

In the next article we will build a speech recognition interface for the OWI 007 Robotic Arm Trainer.

A kit for this project is available from:

#### **Images SI Inc.**

39 Seneca Loop  
Staten Island NY 10314  
718-698-8305 Telephone  
718-982-6145 Fax

SR-07 kit: \$100.00

plus \$8.50 insurance and shipping charge.

To order kit: click [Here](#)