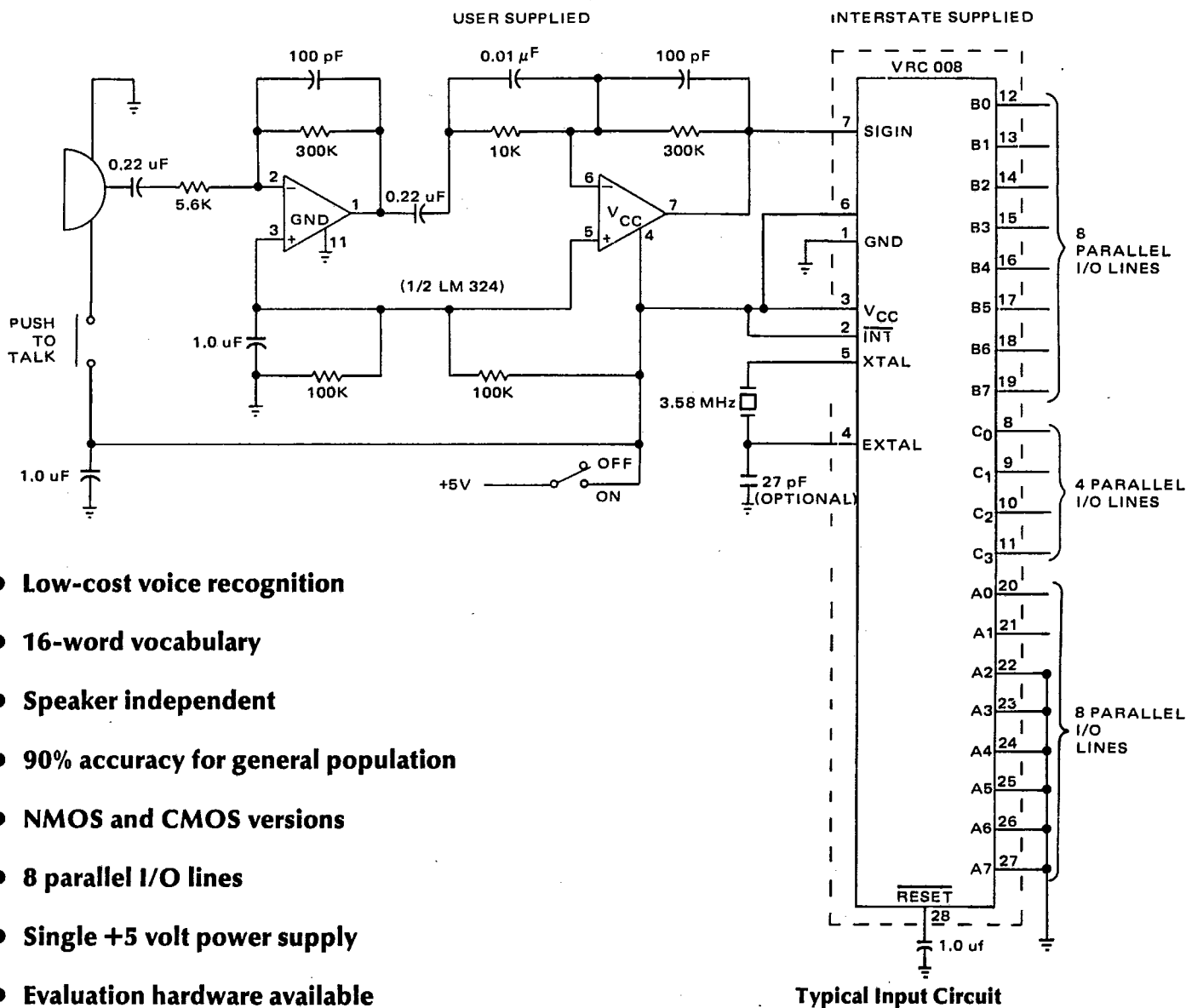




INTERSTATE VOICE PRODUCTS
A FIGGIE INTERNATIONAL COMPANY

VOICE RECOGNITION CHIP

Model VRC008



- Low-cost voice recognition
- 16-word vocabulary
- Speaker independent
- 90% accuracy for general population
- NMOS and CMOS versions
- 8 parallel I/O lines
- Single +5 volt power supply
- Evaluation hardware available

A Low-cost 28-Pin, Single-Chip Voice Recognition System

Interstate's 28-pin, single-chip VRC008 system employs a unique method for processing of analog speech data and recognition of spoken utterances.

Designed for a wide variety of high-volume consumer applications, this microcomputer provides low-cost voice control capability for appliances, toys, games, and other voice automation products. The system is speaker-independent and recognizes, with high accuracy, eight spoken words or phrases, translating verbal commands (i.e., "walk," "stop,"

"channel four," "turn right," etc.) into action via associated circuitry. In a typical application, "wake up" activates the system into a receptive mode and prepares it to accept input speech; the word "relax" stops the system.

Programmable for a selected vocabulary, the VRC008 recognizes speech by detecting significant parameters in the incoming word or phrase and comparing them with the stored sequence of a prespecified vocabulary. With recognition accomplished, the system then outputs a bit pattern for the word number identified. The state sequence and recognition parameters are stored in the on-chip ROM.

Interstate customizes the VRC008 to specific user vocabularies. In this process the customer defines the particular functions to be performed by his product and IEC provides assistance in selecting a vocabulary suited to those functions. The user vocabulary is divided into subsets called syntax nodes, which provide recognition structure to the user format. Therefore, when a specific syntax node is selected, only words contained in the node are candidates for recognition.

MODEL VRC008 SPECIFICATIONS

Performance

Vocabulary size: 16 words or phrases, 8 maximum per node.

Recognition accuracy: 90 percent.

Longest utterance duration: 2.0 seconds.

Minimum word length: 180 milliseconds.

Approximate response time: 200 milliseconds.

Environmental

Operating temperature: 0 to 70°C.

Storage temperature: 55 to 150°C.

Electrical

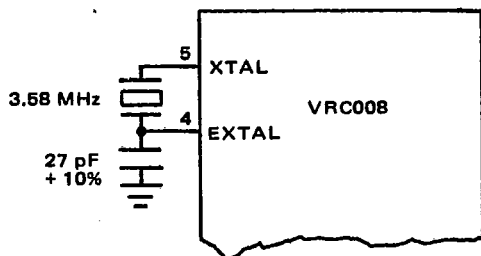
Supply voltage: +5 Vdc ±0.5V.

Audio Input: 0.0 to 5.0 volts maximum.

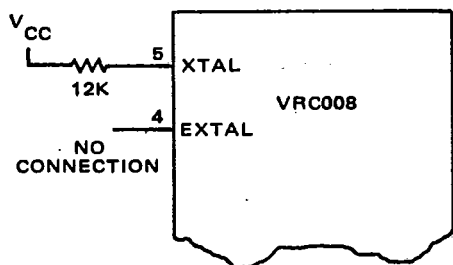
Typical Power Dissipation: 350 mW (NMOS), 20 mW (CMOS).

Internal Clock Generator Options:

A. CRYSTAL OPTION



B. RESISTOR OPTION



APPROXIMATELY 10% ACCURACY EXTERNAL RESISTOR (EXCLUDES RESISTOR TOLERANCE) IF THIS ARRANGEMENT IS USED, RECOGNITION ACCURACY MAY BE DEGRADED.

Digital Input/Output

20 TTL lines configured as a combination of input/output or control lines according to customer specification.

Input/Output Signals

V_{CC} and GND: Power is supplied to the VRC008 using these two pins. V_{CC} is the positive supply voltage (+5V) and GND is the ground connection.

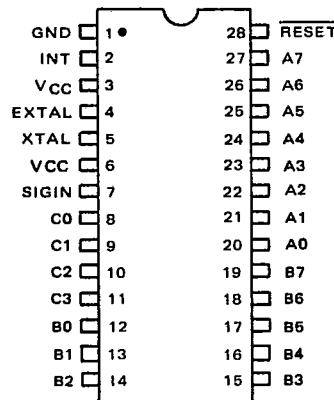
INT: This pin allows an external event to asynchronously interrupt the VRC008. It is normally connected to ground if not specified in user's design.

XTAL and EXTAL: These pins provide connections to the on-chip clock oscillator circuit. A crystal or a resistor is connected to these pins to provide a system clock source with various stability/cost tradeoffs. Lead lengths and stray capacitance on these two pins should be minimized.

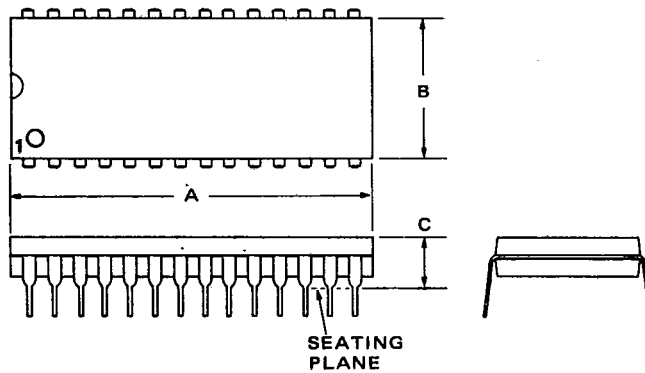
SIGIN: The voice input is applied to this pin. The input signal level should be approximately 1 volt p-p for ambient conditions.

RESET: This pin has a Schmitt Trigger input and an on-chip pullup. The VRC008 can be reset by pulling RESET low.

Input/Output Lines (A0-A7, B0-B7, C0-C3): These lines are arranged into two 8-bit ports (A and B) and a 4-bit port (C). The ports may be configured as inputs, outputs, or control lines as required by customer application.



Pin Assignments



DIM	MILLIMETERS		INCHES	
	MIN	MAX	MIN	MAX
A	36.45	37.21	1.435	1.465
B	13.72	14.22	0.540	0.560
C	4.57	5.08	0.180	0.200

Package Dimensions

EVL008 EVALUATION CHIP

The EVL008 is an EPROM version of the VRC008 with a pre-specified vocabulary. This chip enables customers to become familiar with the speech recognition product. The device demonstrates the versatility of the VRC008 by providing several different vocabulary subsets, selectable by fixed voltage levels on the chip address lines (A0 and A1). The output is provided via the parallel output lines (B0 through B3). Table 1 lists the vocabulary and defines the address, strobe, and output requirements. Figure 1 is a schematic diagram of the test configuration for the EVL008 chip.

Reject

A signal level indicates no match between the word spoken and the stored patterns. Reject is active on each node at output B7. A 1 indicates a reject; a 0 indicates no reject.

Strobe

A 5- μ sec strobe signal is output for any recognized voice input. The strobe will drive a 1 TTL load. A positive strobe is available on pin 8 (C0), a negative strobe on pin 9 (C1).

Vocabulary Item Number

The vocabulary item number is output on the B0 through B3 pins and may sink up to 10 milliamps.

OPERATION

Four different nodes are provided with the EVL008. Node 0 is null, in which case no recognition occurs. The other nodes are listed in table 1.

The input circuit (microphone and amplifier) is designed to transform speech signals into a Bivariate code, representative to some degree of the input and electrically compatible with the microcomputer chip. Hence, strict adherence to the suggested input circuit should be observed. Also, the microphone is normally used no more than a few inches from the mouth. Accuracy is enhanced by clearly pronouncing any speech.

A switched microphone is suggested where extraneous background noise levels are high enough to trigger an otherwise continuously listening system. Background noise and irregularities in speech input can cause performance degradation.

Continued use and familiarity with the system tends to improve response accuracy. The system dynamically responds to differences in word lengths, which can be observed by consciously varying the rate of the utterance.

Table 1. EVL008 Vocabulary

Node		Selection	Vocabulary Word	Item Number			
A1	A0			B3	B2	B1	B0
0	0	Node 0	Null Mode (EVL008 will not react to spoken inputs)	X	X	X	X
0	1	Node 1	Stop	0	0	0	1
			Go Ahead	0	0	1	0
			Search	0	0	1	1
			Faster	0	1	0	0
			Slower	0	1	0	1
			Report	0	1	1	0
1	0	Node 2	Alarm	0	1	1	1
			Lights	1	0	0	0
			Sensors	1	0	0	1
1	1	Node 3	True	1	0	1	0
			False	1	0	1	1
			Repeat	1	1	0	0

Note: Logic 1 = 4 volts
Logic 0 = 0 volts

3

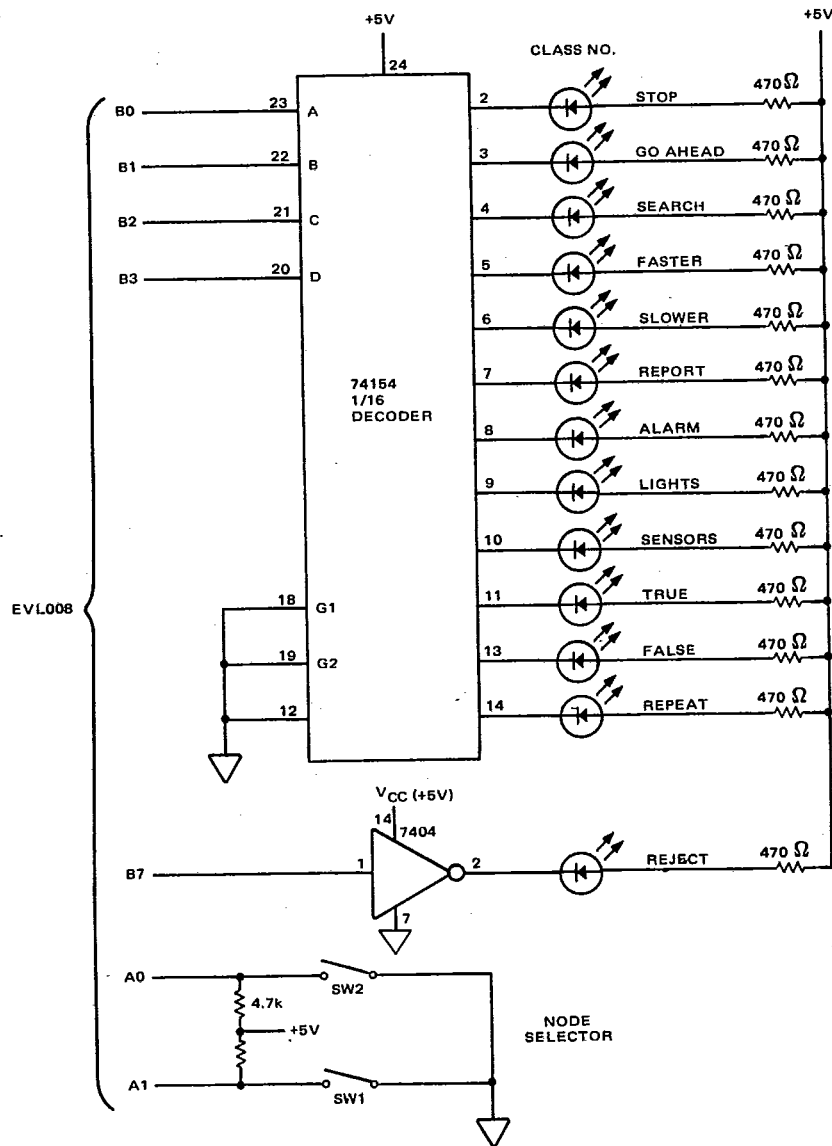


Figure 1. EVL008 Test Configuration

INTERSTATE VOICE PRODUCTS
A FIGGIE INTERNATIONAL COMPANY

Eastern Region Office

1735 Jefferson Davis Highway, Suite 705, Arlington, Virginia 22202-3483
Telephone 703/892-1400

Interstate Voice Products

1849 W. Sequoia Avenue, Orange, California 92668
Telephone 714/937-9010 TWX 910-591-1197 Telex 47-22046 or 65-5419

4