

Votrax®

A Division of Federal Screw Works
500 Stephenson Highway
Troy, Michigan 48084

SC-01 SPEECH SYNTHESIZER

DATA SHEET

Votrax® CMOS Phoneme Speech Synthesizer

GENERAL DESCRIPTION

The SC-01 Speech Synthesizer is a completely self-contained solid state device. This single chip phonetically synthesizes continuous speech, of unlimited vocabulary, from low data rate inputs. Figure 1.

Speech is synthesized by combining phonemes (the building blocks of speech) in the appropriate sequence. The SC-01 Speech Synthesizer contains 64 different phonemes which are accessed by a 6-bit code. It is the proper sequential combination of these phoneme codes that creates continuous speech.

The SC-01 Speech Synthesizer is cost-effective, consumes minimal power and enables in-house product development without vendor dependency. Signals from the SC-01 are applied to an audio output device to amplify and distribute the synthesized speech. See Figure 2.

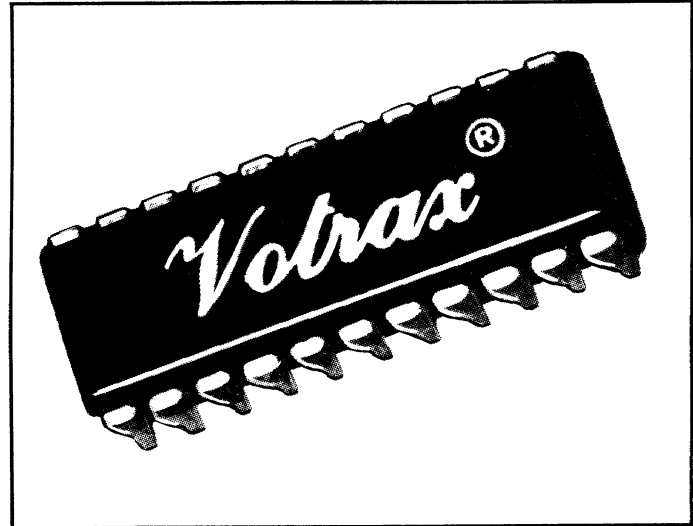


Figure 1. Votrax® SC-01 Speech Synthesizer

FEATURES

- Single CMOS chip
- 70 bits per second
- 22 pin package
- 9 ma. current drain
- Wide voltage supply range
- Latched 5v. compatible inputs
- Digital pitch level inputs
- Automatic inflection
- On-chip master clock circuit
- Optional external master clock
- Variety of voice effects
- Sound effects
- Customer product security

The design of the equipment specified herein is proprietary. Rights for the reproduction and distribution of the data contained herein are granted except for the manufacture and reproduction of the subject equipment.

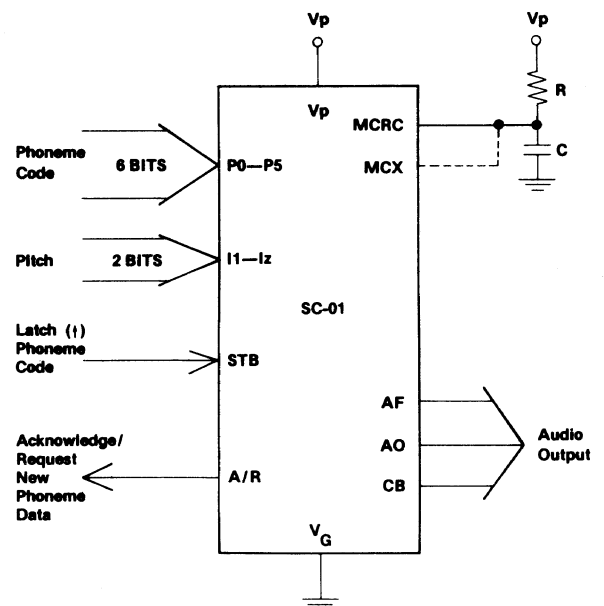


Figure 2. SC-01 Flow Diagram

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

PHYSICAL DESCRIPTION

The SC-01 Speech Synthesizer comprises a 22 pin monolithic (single level or substrate) integrated circuit of CMOS (Complementary Metal On Silicon) design. See Figure 3. The P and N areas of the substrate, made by ion implantation, create a push-pull (Class B) transistor amplification and digital signal processing network. High impedance, rapid signal processing, and minimal current drain result from this technology.

ELECTRICAL DESCRIPTION

The SC-01 Speech Synthesizer is a program-compatible with existing Votrax[®] phoneme synthesizers. It requires 70 bits of data per second for continuous speech production. The 6-bit phoneme codes are 5 volt compatible and are latched for data bus line applications. A phoneme-construction algorithm and filters, within the chip, create the synthesized audio output.

PHONEME DESCRIPTION

Table 1 lists the 64 phonemes produced by the SC-01. Each phoneme code is accompanied by its symbol, average duration time, and an example. The underlined segments of the example word demonstrate the phoneme use, i.e., sound to be pronounced.

Table 2 subdivides the 64 phoneme symbols into seven categories. Each category represents a different production feature. The first six categories are characterized by voiced, fricative (expired voice), and nasal sounds. The seventh category is characterized by phonemes with no sound output.

PHONEME PROGRAMMING

Manual Operations: Votrax[®] maintains a library of phonetically programmed words. Reference to this library and programming manuals will aid in word synthesis.

Automatic Operations: Votrax[®] can supply a micro-computer system for automatic conversion of English text into phoneme sequences. This system is particularly useful for in-house vocabulary development and product security. Contact Votrax[®] for further information.

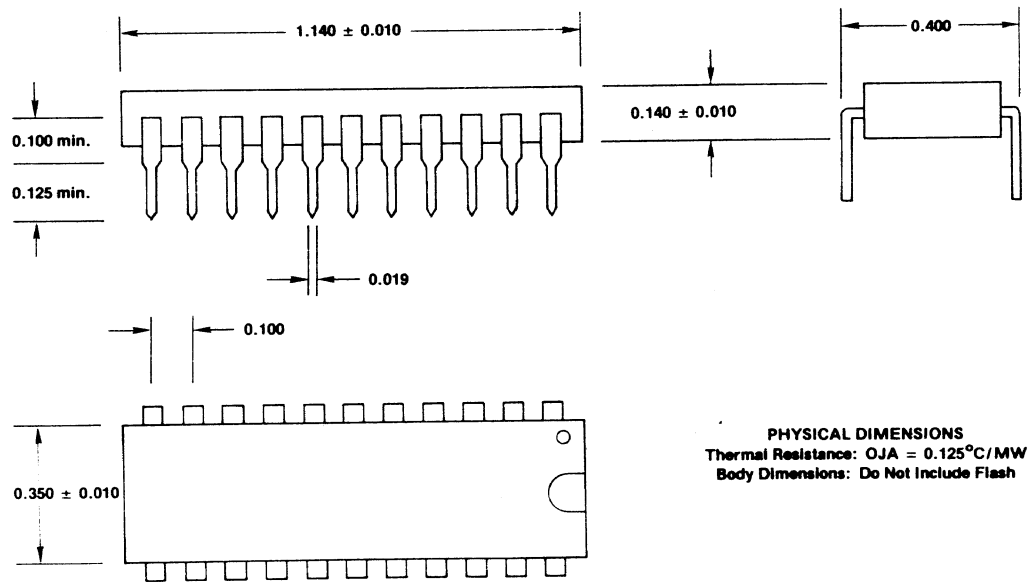
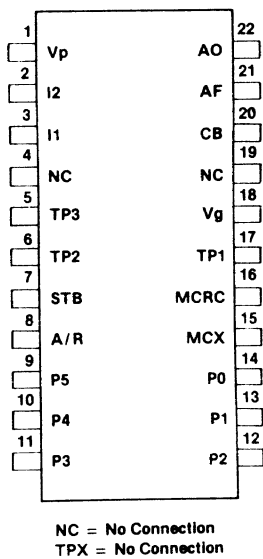


Figure 3. SC-01 Footprint and Outline Dimensions

Votrax[®] reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

Table 1. Phoneme Chart

Phoneme Code	Phoneme Symbol	Duration (ms)	Example Word
ØØ	EH3	59	jack <u>e</u> t
Ø1	EH2	71	enl <u>e</u> st
Ø2	EH1	121	heav <u>y</u>
Ø3	PAØ	47	no sound
Ø4	DT	47	b <u>u</u> tter
Ø5	A2	71	mad <u>e</u>
Ø6	A1	103	mad <u>e</u>
Ø7	ZH	90	azur <u>e</u>
Ø8	AH2	71	hon <u>e</u> st
Ø9	I3	55	inhib <u>i</u> t
ØA	I2	80	inhib <u>i</u> t
ØB	I1	121	inhib <u>i</u> t
ØC	M	103	mat <u>h</u>
ØD	N	80	sun <u>n</u>
ØE	B	71	bag <u>g</u>
ØF	V	71	van <u>n</u>
1Ø	CH*	71	ch <u>i</u> p
11	SH	121	sh <u>o</u> p
12	Z	71	zoo <u>z</u>
13	AW1	146	lawf <u>u</u> l
14	NG	121	th <u>i</u> ng
15	AH1	146	fat <u>h</u> er
16	OO1	103	look <u>i</u> ng
17	OO	185	book <u>o</u>
18	L	103	land <u>l</u>
19	K	80	tr <u>i</u> ck
1A	J*	47	judg <u>e</u>
1B	H	71	hell <u>o</u>
1C	G	71	g <u>e</u> t
1D	F	103	fast <u>f</u>
1E	D	55	paid <u>d</u>
1F	S	90	pass <u>s</u>

Phoneme Code	Phoneme Symbol	Duration (ms)	Example Word
2Ø	A	185	day <u>a</u>
21	AY	65	day <u>a</u>
22	Y1	80	yard <u>y</u>
23	UH3	47	miss <u>i</u> on
24	AH	250	m <u>o</u> p
25	P	103	past <u>p</u>
26	O	185	c <u>o</u> ld
27	I	185	pin <u>i</u>
28	U	185	mov <u>u</u> e
29	Y	103	an <u>y</u>
2A	T	71	tap <u>t</u>
2B	R	90	red <u>r</u>
2C	E	185	me <u>e</u> t
2D	W	80	win <u>w</u>
2E	AE	185	dad <u>a</u>
2F	AE1	103	af <u>a</u> ter
3Ø	AW2	90	salt <u>y</u>
31	UH2	71	ab <u>o</u> ut
32	UH1	103	unc <u>l</u> e
33	UH	185	cup <u>u</u>
34	O2	80	for <u>o</u>
35	O1	121	ab <u>o</u> ard
36	IU	59	you <u>u</u>
37	U1	90	you <u>u</u>
38	THV	80	th <u>e</u>
39	TH	71	th <u>i</u> n
3A	ER	146	bird <u>r</u>
3B	EH	185	g <u>e</u> t
3C	E1	121	be <u>e</u>
3D	AW	250	call <u>a</u>
3E	PA1	185	no sound
3F	STOP	47	no sound

* T must precede CH to produce J sound.
D must precede J to produce CH sound.

Table 2. Phoneme Categories According to Production Features

Voiced					'Voiced' Fricat.	'Voiced' Stop	Fricat. Stop	Frica-tive	Nasal	No Sound
E	EH	AE	UH	OO1	Z	B	T	S	M	PAØ
E1	EH1	AE1	UH1	R	ZH	D	DT	SH	N	PA1
Y	EH2	AH	UH2	ER	J	G	K	CH	NG	STOP
Y1	EH3	AH1	UH3	L	V		P	TH		
I	A	AH2	O	IU	THV			F		
I1	A1	AW	O1	U				H		
I2	A2	AW1	O2	U1						
I3	AY	AW2	OO	W						

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

SIGNAL DESCRIPTION (See Figures 4 and 5)

Phoneme 6-Bit Selection Code (P0-P5): Data input is to six pins. Latching is controlled by the strobe (STB) signal.

Strobe (STB): Latching occurs on rising edge of strobe signal.

Inflection Level Setting (I1, I2): Instantaneously set pitch of voiced phonemes.

Acknowledge/Request (A/R): Acknowledges receipt of phoneme data (signal goes from high to low one master clock cycle following active edge of STB signal). Also indicates timing out of old phoneme concurrent with request for new phoneme data (signal goes from low to high).

NOTE

If external phoneme timing is desired, phoneme requests can be ignored. However, best speech is realized with internal timing.

Master Clock Resistor-Capacitor (MCRC): This input determines the internal master clock frequency. Select R-C values for 720 KHz to achieve standard phoneme timing. Connect this input to MCX when using internal clock; ground when using external clock.

NOTE

Varying clock frequency varies voice and sound effects. As clock frequency decreases, audio frequency decreases and phoneme timing lengthens. Figures 6 and 7 illustrate manual and DAC (Digital to Analog Converter) voice variation schematics, respectively.

Master Clock External (MCX): Allows control by an external clock signal.

NOTE

Ground MCRC during MCX operation.

Audio Output (AO): Supplies analog signal to audio output device.

Audio Feedback (AF): Used with Class A or Class B transistor audio amplifiers for added stability.

Class B (CB): Current source for Class B transistor audio amplifier.

Table 3. Timing Specifications

CHARACTERISTIC	SYMBOL	MIN	TYP	MAX	UNIT
Input Setup Time (P _i to STB)	T _S			450	NS
Input Hold Time (P _i to STB)	T _H			0	NS
Rise Time of STB Edge (.8v to 4v)	T _{RS}			100	NS
A/R Width (A/R Connected to STB) +	T _{ARW}	1	1.3	2	μs
STB Width	T _{SW}	200			NS
STB Low*	T _{SL}	*			
Propagation Delay (STB to A/R after 2μs)	T _{DAR}			500	NS
A/R Rise Time (Capacitive load = 30pf)	T _{RAR}			100	NS
A/R Fall Time (Capacitive load = 30pf)	T _{FAR}			100	NS
Time from A/R Request to STB Service)	T _{ARS}	0	500		μs
Time of Phoneme Duration +	T _{PH}	47	107	250	MS

+ Dependent on Master Clock frequency: 720KHz = 1.3μs

* Strobe must remain low (64 x master clock period) before rising edge.

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

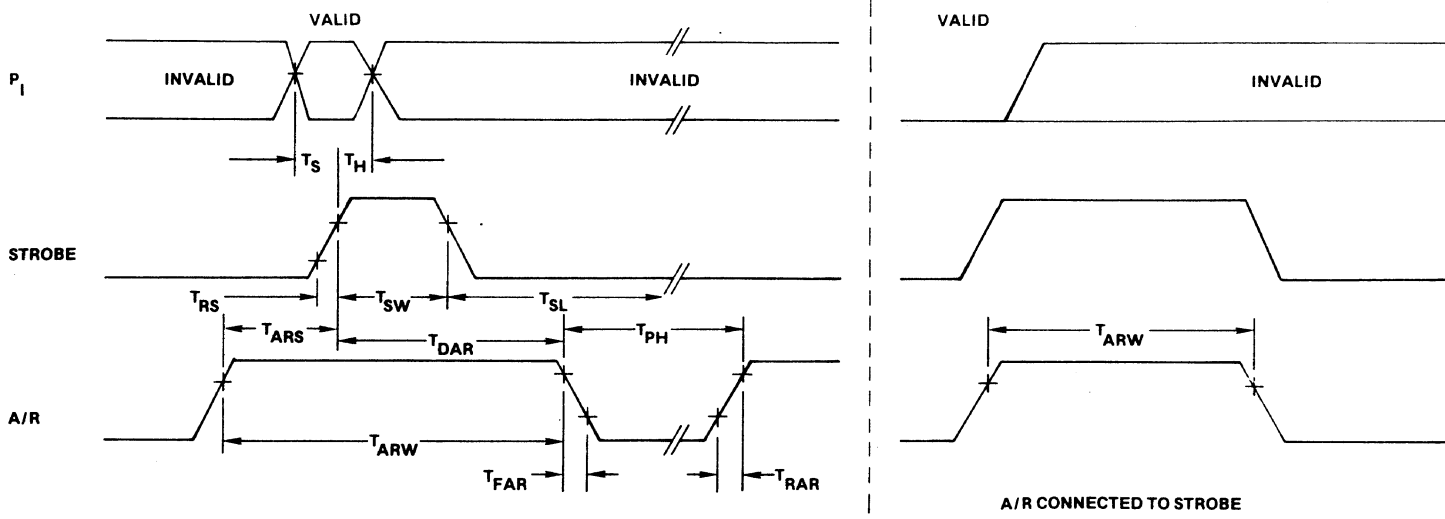


Figure 4. Timing Diagram

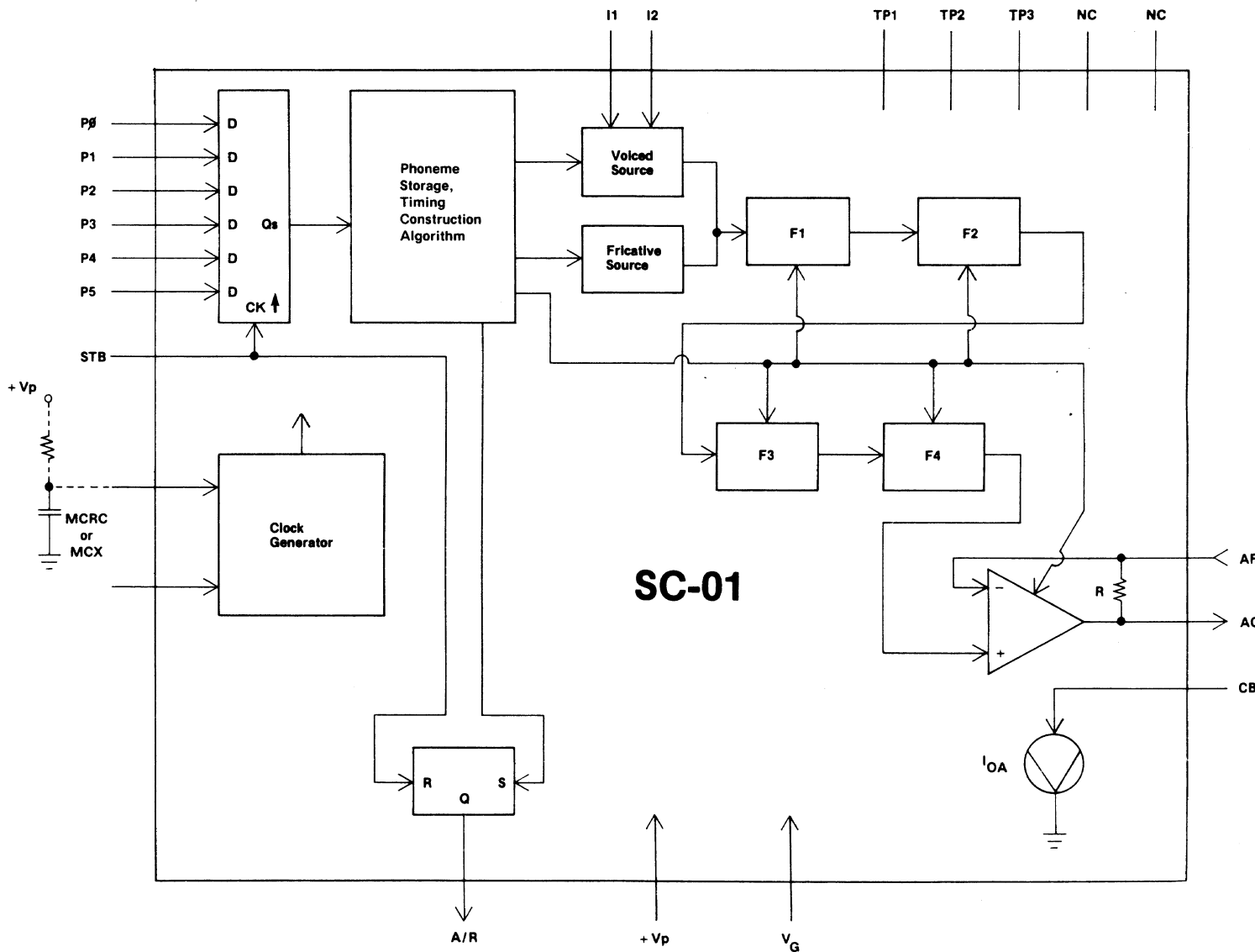


Figure 5. SC-01 Block Diagram

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

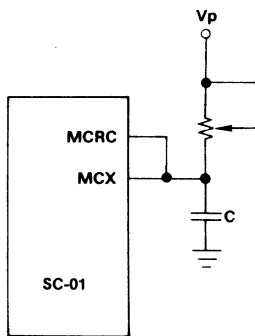


Figure 6. Variable Voice by Potentiometer Control

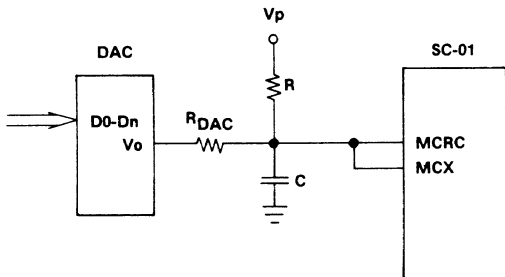


Figure 7. Variable Voice by DAC Current Injection

TYPICAL APPLICATIONS

General: The SC-01 Speech Synthesizer is easily designed into systems ranging in complexity from ROM/counters to microprocessor controllers.

Single Message System: See Figure 8. When the counter is released (START is TRUE), the message is clocked out of the ROM by the A/R signal. The system must be stopped when DONE is TRUE.

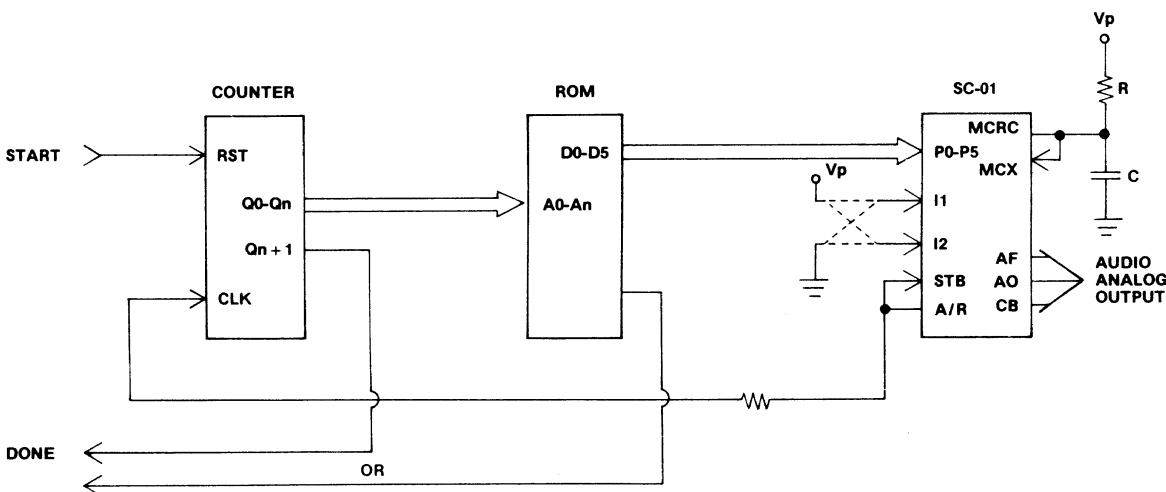


Figure 8. Single Message System

NOTE

Data at address 0 must be a pause phoneme code.

Multiple Message, Fixed Block Size: See Figure 9. Message address block is loaded into the counter. The message is then clocked out of the ROM by the A/R signal.

NOTE

Message Block = 2^n maximum.

Multiple Message, Variable Block Size: See Figure 10. The microprocessor loads phonemes into a data bus. The A/R signal generates an interrupt request for each new phoneme.

CONNECTING THE AUDIO OUTPUT DEVICE

Audio Output: The AO signal has a maximum voltage swing of 1.5 volts below V_p or above V_G , depending upon the phoneme selected. Furthermore, the AO signal is D.C. biased by a factor of $V_p/2$.

NOTE

If additional audio amplifier stability is not needed, connect the AO pin to the AF pin.

Class A Amplifier: See Figure 11. For a single transistor amplifier, the selection of R, C, or R_s values depends upon the value of V_p and the desired audio level.

NOTE

The CB pin is not used for Class A amplifiers.

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

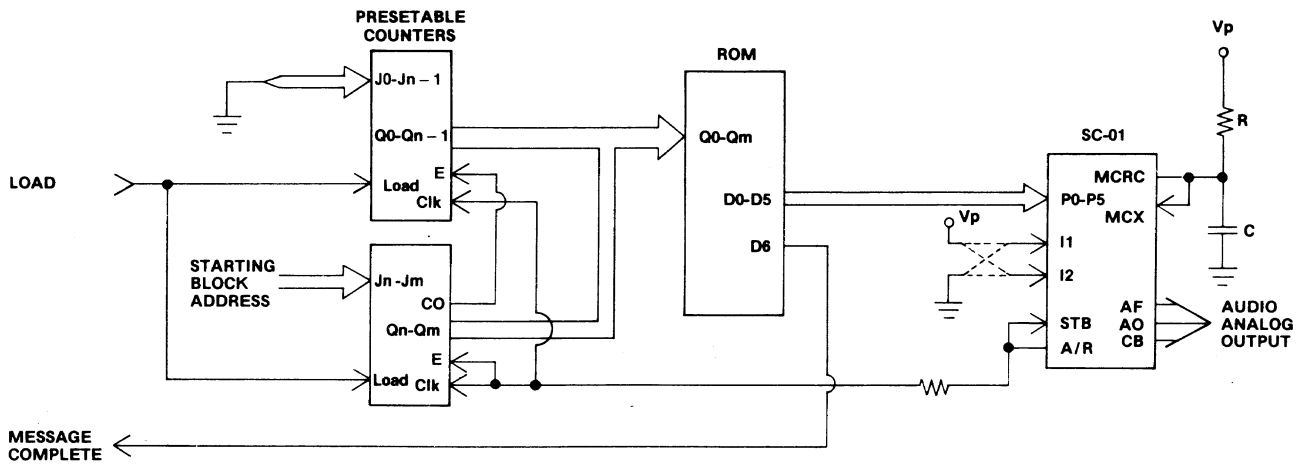


Figure 9. Multiple Message, Fixed Block Size

Class B Amplifier: See Figure 12. A current source (CB) is required for this push-pull amplifier.

NOTE

Minimum power is consumed when speech is inactive. When $V_p = +12.0$ volts and $R_s = 40$ ohms, the bias current drain is approximately 3.5 milliamps.

Controlling Audio Output Power: See Figure 13. A resistor or potentiometer from the speaker to ground can be used to control the audio output power.

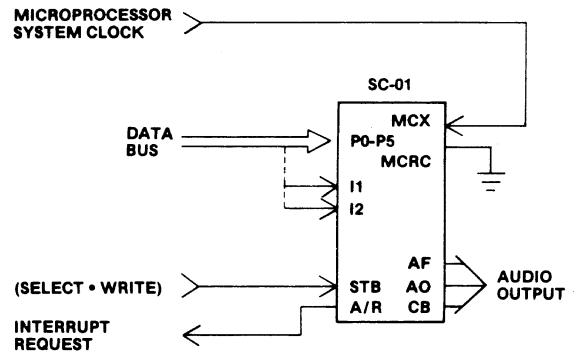


Figure 10. Multiple Message, Variable Block Size

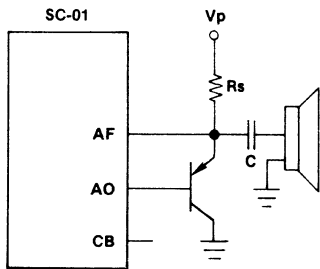


Figure 11. Class A Amplifier

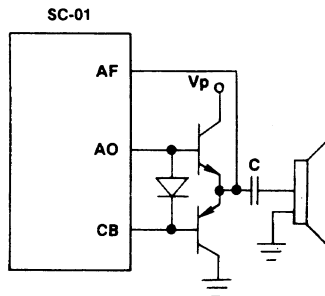


Figure 12. Class B Amplifier

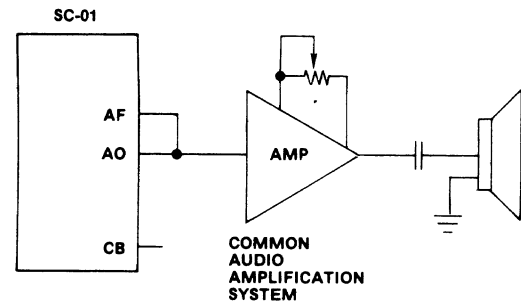


Figure 13. Controlling Audio Output Power

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

