votrax®

VOTRAX INTERNATIONAL, INC.
1394 Rankin
Troy, Michigan 48083-4074
(313) 588-2050 TWX-8102324140
Votrax-TRMI

"Technology that Talks"

VOTRAX SC-01 Phoneme Speech Synthesizer

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.

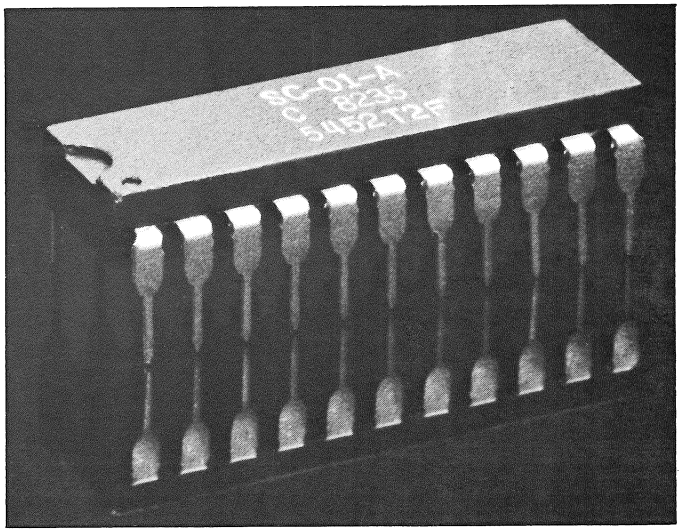


Figure 1. Votrax® SC-01 Speech Synthesizer

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

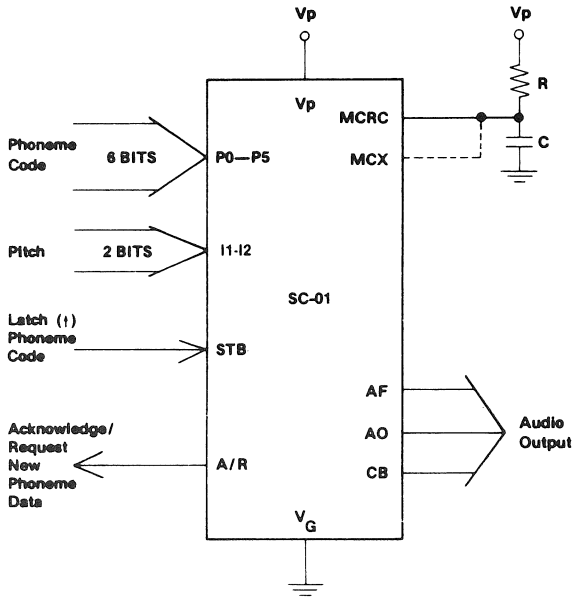


Figure 2. SC-01 Flow Diagram

PHYSICAL DESCRIPTION

The SC-01 Speech Synthesizer is a 22 pin Large Scale Integrated Circuit which contains all the circuitry necessary to generate phonetically synthesized speech. The SC-01 is fabricated using CMOS technology, which offers high input impedance and low power drain.

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 logic compatible and are latched for data bus 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.

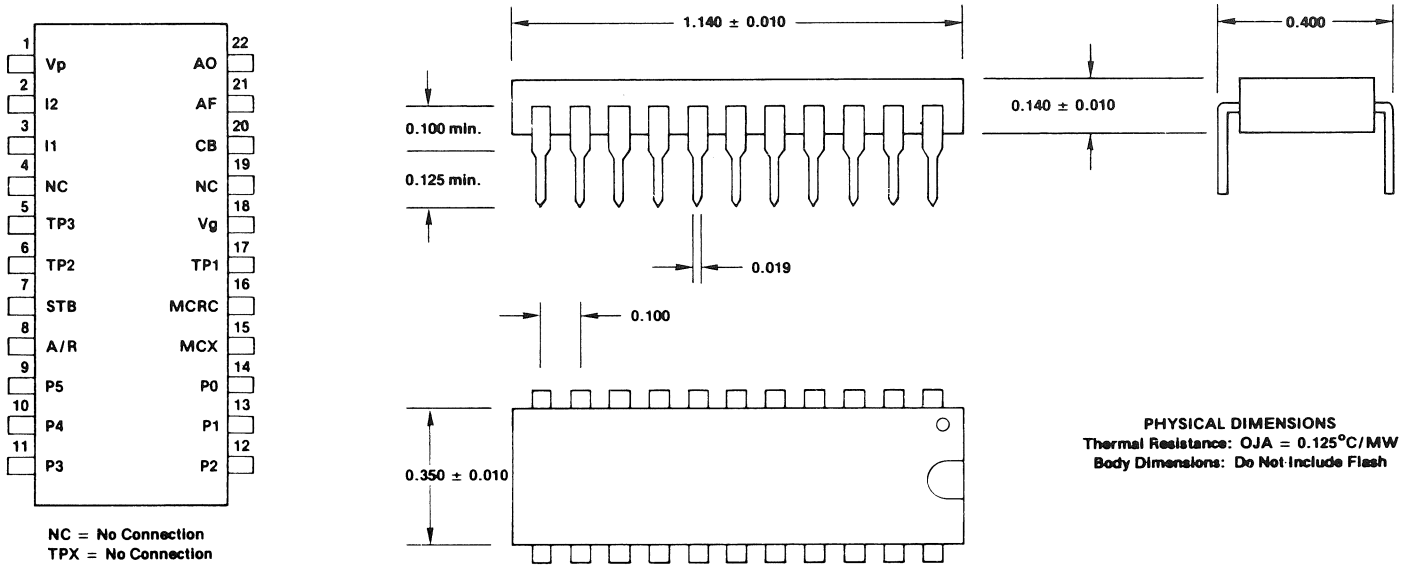


Figure 3. SC-01 Footprint and Outline Dimensions

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	jacket
Ø1	EH2	71	enlist
Ø2	EH1	121	heavy
Ø3	PAØ	47	no sound
Ø4	DT	47	butter
Ø5	A2	71	made
Ø6	A1	103	made
Ø7	ZH	90	azure
Ø8	AH2	71	honest
Ø9	I3	55	inhibit
ØA	I2	80	inhibit
ØB	I1	121	inhibit
ØC	M	103	mat
ØD	N	80	sun
ØE	B	71	bag
ØF	V	71	van
1Ø	CH*	71	chip
11	SH	121	shop
12	Z	71	zoo
13	AW1	146	lawful
14	NG	121	thing
15	AH1	146	father
16	OO1	103	looking
17	OO	185	book
18	L	103	land
19	K	80	trick
1A	J*	47	judge
1B	H	71	hello
1C	G	71	get
1D	F	103	fast
1E	D	55	paid
1F	S	90	pass

Phoneme Code	Phoneme Symbol	Duration (ms)	Example Word
2Ø	A	185	day
21	AY	65	day
22	Y1	80	yard
23	UH3	47	mission
24	AH	250	mop
25	P	103	past
26	O	185	cold
27	I	185	pin
28	U	185	move
29	Y	103	any
2A	T	71	tap
2B	R	90	red
2C	E	185	meet
2D	W	80	win
2E	AE	185	dad
2F	AE1	103	after
3Ø	AW2	90	salty
31	UH2	71	about
32	UH1	103	uncle
33	UH	185	cup
34	O2	80	for
35	O1	121	aboard
36	IU	59	you
37	U1	90	you
38	THV	80	the
39	TH	71	thin
3A	ER	146	bird
3B	EH	185	get
3C	E1	121	be
3D	AW	250	call
3E	PA1	185	no sound
3F	STOP	47	no sound

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

Table 2. Phoneme Categories According to Production Features

Voiced					'Voiced' Fricat.	'Voiced' Stop	Fricative Stop	Fricative	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 (PØ-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 sets pitch level of voiced phonemes.

Acknowledge/Request ($\bar{A}$/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	Ø			NS
Rise Time of STB Edge (.8V to 4V)	T_{RS}			1ØØ	NS
A/R Width ($\bar{A}$/R Connected to STB) +	T_{ARW}	1	1.3	2	μs
STB Width	T_{SW}	2ØØ			NS
STB Low*	T_{SL}	*			NS
Propagation Delay (STB to A/R after T_{ARW})	T_{DAR}			5ØØ	NS
A/R Rise Time (Capacitive load = 3Øpf)	T_{RAR}			1ØØ	NS
A/R Fall Time (Capacitive load = 3Øpf)	T_{FAR}			1ØØ	NS
Time from $\bar{A}$/R Request to STB Service)	T_{ARS}	Ø		500	μs
Time of Phoneme Duration +	T_{PH}	47	1Ø7	25Ø	MS

+ Dependent on Master Clock frequency: 720kHz

* Strobe must remain low (72x 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.

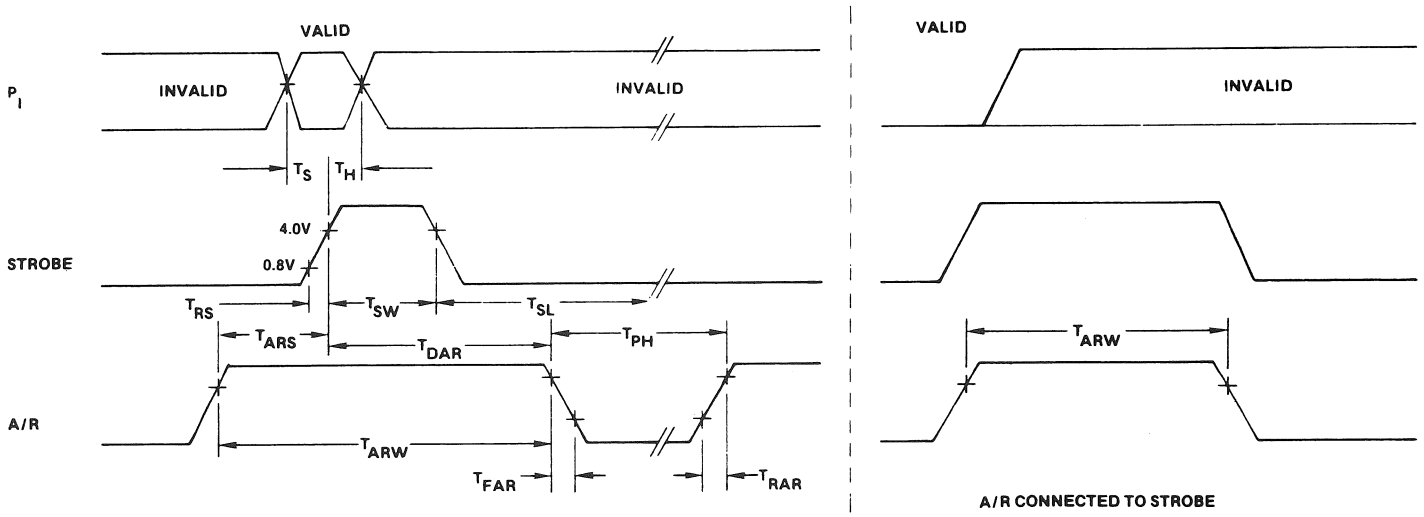


Figure 4. Timing Diagram

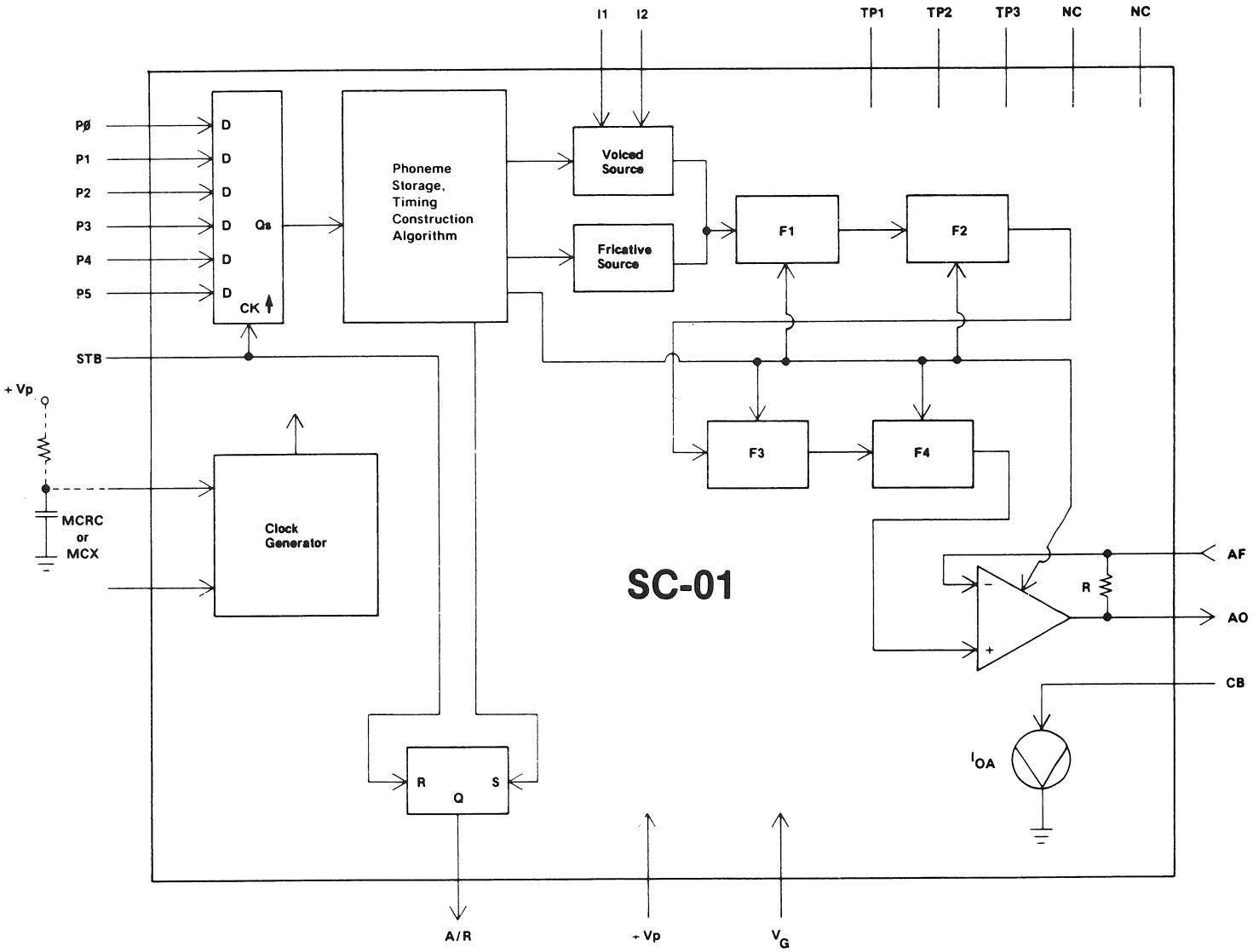


Figure 5. SC-01 Block Diagram

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

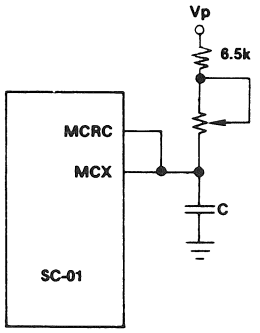


Figure 6. Variable Voice by Potentiometer Control

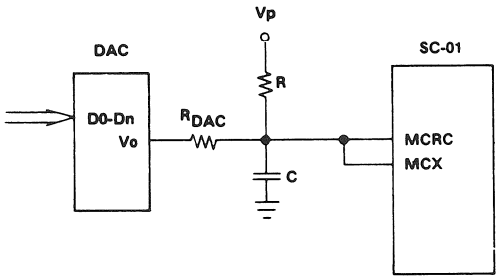


Figure 7. Variable Voice by DAC Current Injection

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. Note: When using A/R tied to STB, connect a .01 uf capacitor to TP3 to insure power up reset of SC-01.

NOTE

Data at address Ø 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 peak to peak voltage swing of .26 times Vp, depending upon the phoneme selected, and the AO signal is D.C. biased.

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 Vp and the desired audio level.

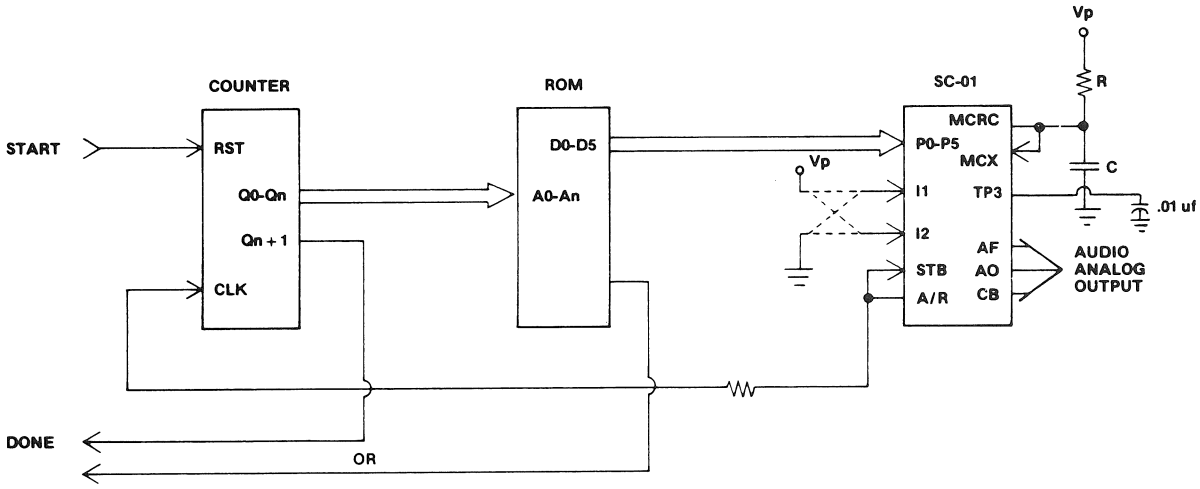


Figure 8. Single Message System

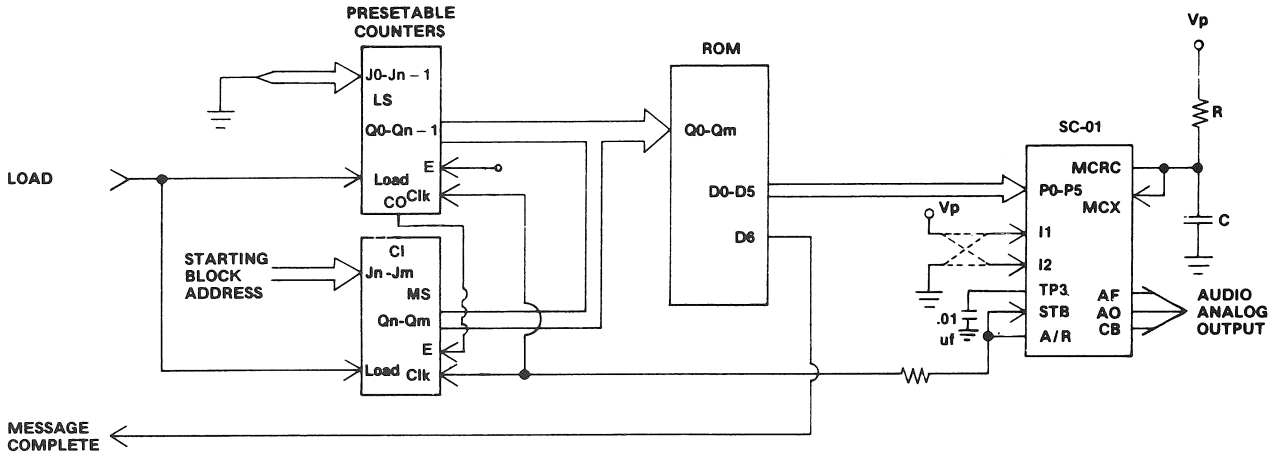


Figure 9. Multiple Message, Fixed Block Size

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 Vp = +12.Ø volts and R_s = 4Ø 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.

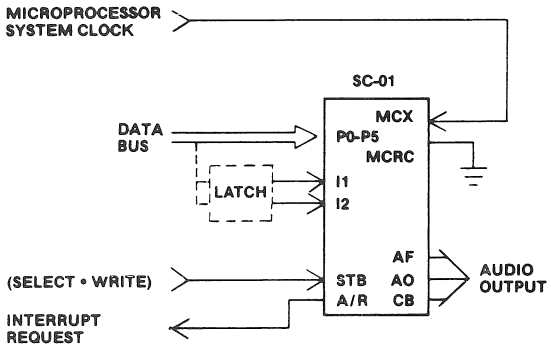


Figure 10. Multiple Message, Variable Block Size

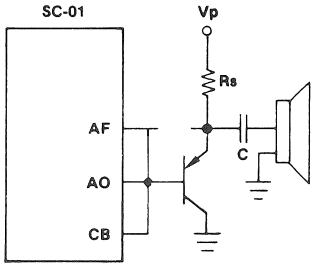


Figure 11. Class A Amplifier

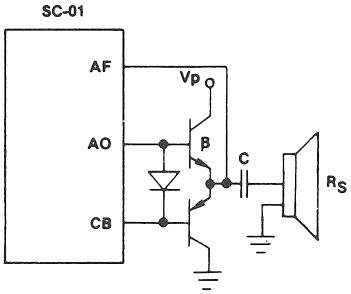


Figure 12. Class B Amplifier*

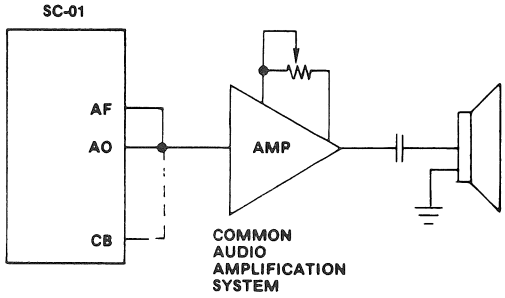


Figure 13. Controlling Audio Output Power

*For Class B Amplifier: (β) x (R_S min.) = 81.6 x (Vp) where β is beta or current gain of transistor. The AO line is protected by an internal series current limiting resistor of 90 ohms maximum. If more current is required of the SC-01, then the above formula indicates distortion will occur.

Table 4. Analog Output Specifications

CHARACTERISTIC	MIN	MAX	UNIT
Output Voltage (AH Phoneme)	.18 x Vp	.26 x Vp	Vp-p
Output Bias Current ** (.6V< CB< Vp)	3.5	7.3	mA

ELECTRICAL CHARACTERISTICS: T_o = Ø to 7Ø°C, Vp = 10 V_{DC} to 14 V_{DC}

CHARACTERISTIC		MIN	TYP	MAX	UNIT
Digital Input Impedance		1 meg.			Ohm
Input Capacitance (P_I, STB)				3	pf
Input Capacitance (I1, I2, MCX)				8	pf
Digital Input Logic "Ø" (except I1, I2, MCX)		V_G - Ø.5		V_G + Ø.8	V_{DC}
Digital Input Logic "Ø" (MCX)				V_G + 1.Ø	V_{DC}
Digital Input Logic "Ø" (I1, I2)				.2 x Vp	V_{DC}
Digital Input Logic "1" (except I1, I2, MCX)		V_G + 4.Ø		Vp + Ø.5	V_{DC}
Digital Input Logic "1" (I1, I2)		.8 x Vp			V_{DC}
Digital Input Logic "1" (MCX)		4.6			V_{DC}
Digital Output Logic "Ø" (I sink = Ø.8mA)				V_G+Ø.5	V_{DC}
Digital Output Logic "1" (I source = Ø.5mA)		Vp–Ø.5			V_{DC}
Power Supply Current	Vp = 9V		9.1		mA
	Vp = 9V**		11	18	mA
	Vp = 14V**		18	27	mA
*Master Clock Frequency			720K		Hz
MCX Input Duty Cycle		60:40		40:60	%
Master Clock Resistor Value (MCRC)***		6.5k			Ohm
Master Clock Capacitor Value (MCRC)***				3ØØ	pf

**Variable*

***With CB, AF, AO connected for Class B audio amplifier (see APPLICATION NOTES)*

****Frequency of Master Clock* ≃ 1.25 / RC

Note: TP1, TP2 must be left open for normal operation.

Table 5. Absolute Maximum Ratings

ABSOLUTE MAXIMUM RATINGS *

RATING	SYMBOL	VALUE	UNIT
Power Supply Voltage	Vp	20	V_{DC}
Power Dissipation at 25°C	P_{DM}	650	mW
Derating Above 25°C		5	mW/°C
Operating Ambient Temperature	T_o	Ø to 7Ø	°C
Storage Temperature	T_{STG}	−55 to 125	°C
Input Voltage	V_{INM}	−Ø.5 to Vp+Ø.5	V_{DC}
DC Current Max. Above Vp+Ø.5V	I_{INM}	1.Ø	ma
Lead Temperature (soldering 1Ø sec.)	T_L	3ØØ	°C

** Operation above these limits could damage the device.*

NORMAL OPERATING CONDITIONS: $10v \leq Vp \leq 14v$, $0° C \leq T_o \leq 70° C$

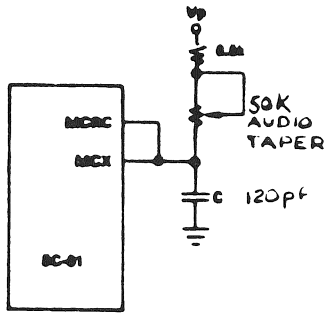


Figure 6. Variable Voice by Potentiometer Control

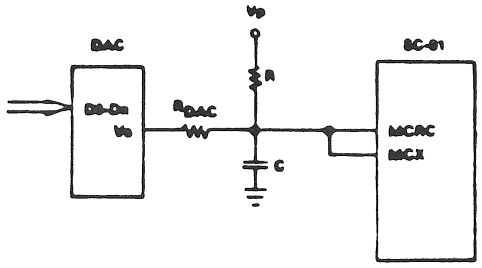
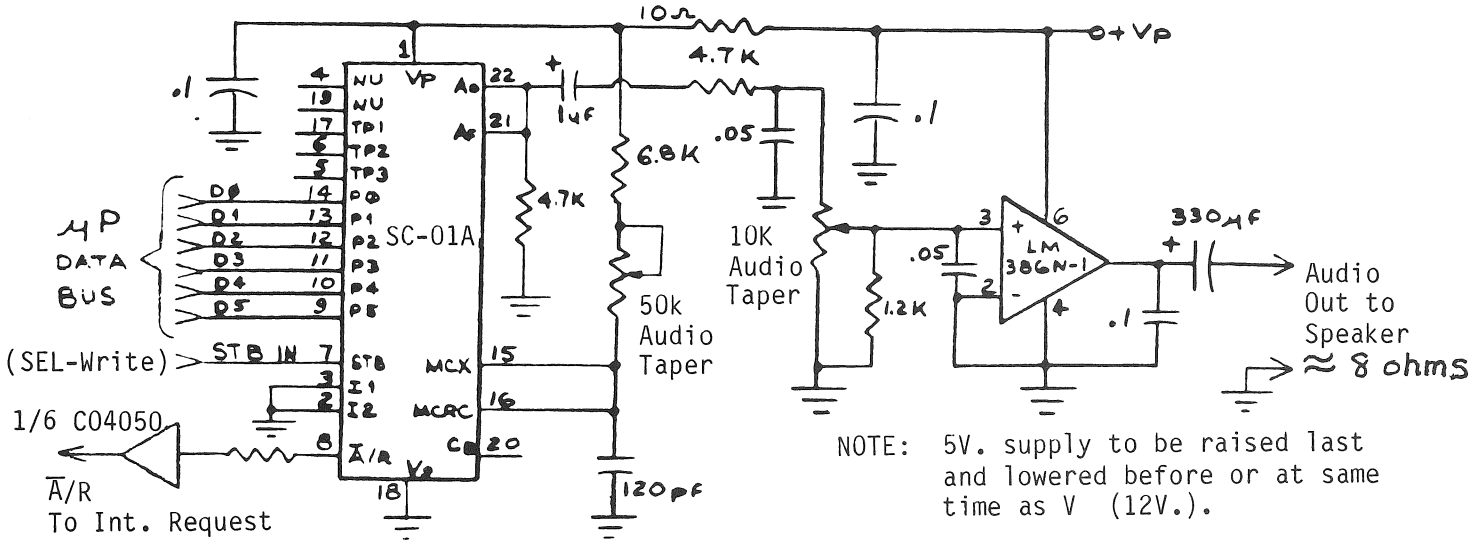


Figure 7. Variable Voice by DAC Current Injection



NOTE: 5V. supply to be raised last and lowered before or at same time as V (12V.).

FIGURE 8. Typical Application

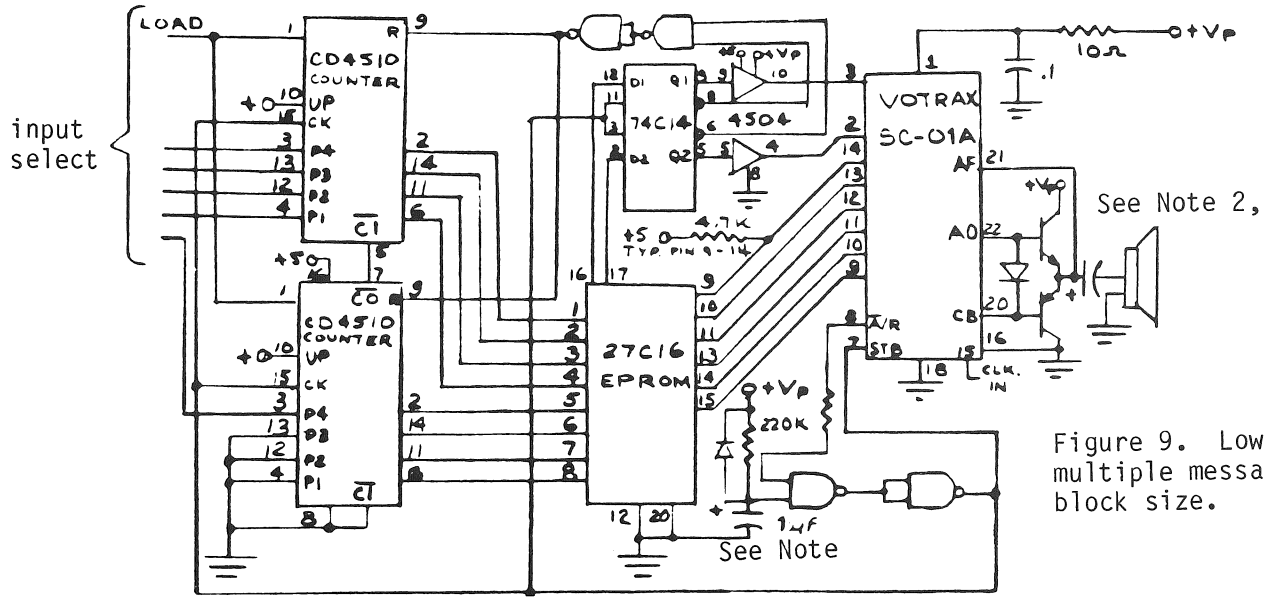


Figure 9. Low power, multiple message, 8 Byte block size.

NOTE*: Circuit uses A/R line as strobe. R-C provides an initial strobe to insure power up reset of SC-01 A/R output.

CIRCUIT DESCRIPTION

The following is a brief functional description of the pin-outs on the SPEECH-PAC™ (see Schematic Figure 2). For a more detailed explanation of their operation and interaction, refer to SYSTEM OPERATION.

EDGE CARD PARALLEL INTERFACE (J1):

PINS*	FUNCTION*
1,2,3,4	Power and Ground connections. Note that power may be supplied through either the Edge Connector (J1) or the Post Connector (J2).
12	SEL: controls the mode of operation (1 or 2) of the SPEECH-PAC™.
15	$\overline{\text{SEL}}$: inverted state of SEL at any time.
6	LOAD: pulse input used during Mode 1 operation to load the desired ROM Address.
7	$\overline{\text{EOM}}$: used during Mode 1 operation to indicate when the last phoneme of the requested word/phrase has been loaded into the SC-01 Speech Chip.
A,B,C,D,E,F	PØ-P5: during Mode 2 operation, these lines are loaded with the desired phoneme code.
A,B,C,D,E,F,H,J	LØ-L7: during Mode 1 operation, these lines are loaded with the ROM Address of the desired word/phrase.
14	L8: this pin is provided to facilitate expansion of the system from a 16K ROM (2716) to a 32K ROM (2532). This expansion is easily accomplished by: a) Breaking the connection between Jumpers 2 and 3. b) Making a connection between Jumpers 1 and 2. c) Replacing the current 2716 ROM with a 2532 ROM.

PINS*	FUNCTION*
14	From this point on, Pin 14 (L8) would be utilized as the MSB of the ROM Address.
K,L	I1, I2: used during either Mode 1 or Mode 2 to set the inflection level of the SC-01 Speech Chip.
9	$\overline{A}$/R: used during Mode 2 operation to determine when the SC-01 Speech Chip is ready for another phoneme.
10	STB: used during Mode 2 operation to load the phoneme present on PØ-P5 into the SC-01 Speech Chip.

* Note: all inputs are one TTL load.

POST CONNECTOR (J2):

PINS	FUNCTION
4,5,6,7	Power and Ground connections. Note that power may be supplied through either the Edge Connector (J1) or the Post Connector (J2).
1	AUDIO OUT: output signal from the Audio Amplifier.
2	AUDIO GND: Ground level for the Audio Amplifier.
8	MCKIN: provided to facilitate the utilization of an external clock by the user. This may be accomplished by: a) Breaking the connection between Jumpers 4 and 5. (Disables the internal clock.) b) Making a connection between Jumpers 5 and 6, and a second connection between Jumpers 4 and 7. From this point on, the clock signal present on Pin 8 (MCKIN) will be used to drive the system.

SYSTEM OPERATION

The following is a description of the two modes of operation available on the SPEECH-PAC:

MODE 1 (WORD MODE)

This mode is utilized to access words or phrases pre-stored in the ROM. The sequence of events involved is as follows:

a) SEL is driven low to select Mode 1. This de-selects the phoneme lines (PØ-P5) by disabling U6 (see Schematic Diagram #1), and enabling EPROM U7.

b) The address of the desired word or phrase is placed on ROM Address load line LØ-L7. Refer to Table 4 for a complete list of the available words/phrases and their corresponding addresses. Note that when loading the address, LØ is the least significant bit.

c) The LOAD line is driven high. This loads the address into the ROM JAM COUNTERS (U2, U3, U4) and thereby to the ROM itself. Note: for proper operation, the data on lines LØ-L7 must be valid for a period of 30 nsecs. prior to the load line going high and while load line is high. Following this, the LOAD line must be driven low -- otherwise the ROM JAM COUNTERS will not be updated, and the SPEECH-PAC™ will remain "stuck" on one phoneme.

d) The phoneme stored in ROM at the presented address is output and loaded into the SC-01 Speech Chip. Datalines D6 and D7 on the ROM (U7) are high, thereby allowing incrementation of the ROM JAM COUNTERS to the next address and setting-up the next phoneme. When the SC-01 (U11) is ready for the next phoneme, it raises $\overline{A}/R$ (Pin 8, U11) high. Since, in this mode, $\overline{A}/R$ is connected to STB (Pin 7, U11), STB is also driven high thereby loading the phoneme into the SC-01. $\overline{A}/R$ then returns low, driving STB back low. It is this pulse which increments the ROM JAM COUNTERS (U2, U3, and U4).

MODE 1 (WORD MODE)

e) This sequence is repeated until the last phoneme of the word/phrase is outputted from the ROM. In this phoneme, Bit #6 (D6, U7) is low. This signal clears the ROM JAM COUNTERS, thereby resetting the address presented to the ROM. This reset address accesses a phoneme in the ROM in which Bit #7 (D7, U7) is low. This signal in effect disables the ROM JAM COUNTERS and drives line $\overline{\text{EOM}}$ low, thereby indicating the end of the sequence to the user. The last phoneme is still being vocalized via the SC-01 at this time. Therefore, if the user wishes to output another word/phrase immediately after completion of the current word/phrase, simply return to Step 'b' immediately and start again. If a pause is desired between the current word/phrase and the next, simply wait the desired length of time and then return to step 'b'.

MODE 2 (PHONEME MODE)

In this mode, the user may construct his own words, phrases, or sound effects by loading phonemes directly to the SC-01 Speech Chip. The sequence of events involved are as follows:

a) SEL line is driven high to select Mode 2 operation. This selects the phoneme lines PØ-P5 by enabling U6. This also disables U7 and the hardware by which the STB signal (Pin 7, U11) is controlled by the $\overline{\text{A}}$/R signal (Pin 8 U11). This permits complete control of the SC-01 Speech Chip by the user.

b) The desired phoneme code is placed on phoneme lines PØ-P5. Refer to Tables 1 & 2 for a complete list of valid phonemes and their corresponding codes. Note that PØ is the least significant Bit when strobing in phoneme codes.

c) A test is then made to determine the state of the $\overline{\text{A}}$/R line. When this line is high, it indicates that the SC-01 Speech Chip is ready to receive a phoneme. Once this condition has been detected, the STB line should be driven high to strobe in the phoneme code present on lines PØ-P5 into the SC-01. The STB line should then be driven low again, after which the $\overline{\text{A}}$/R line will go low until the SC-01 is ready for the next phoneme. This process is repeated until the entire phoneme sequence desired has been output.

NOTE: If desired, the SC-01 can be configured to control itself in terms of loading phonemes. This is easily accomplished by connecting the $\overline{A}/R$ line to the STB line. In this method, the $\overline{A}/R$ line is monitored and, each time this line goes low, the next phoneme is placed on lines PØ-P5. When the SC-01 is ready for the next phoneme, it raises the $\overline{A}/R$ line, thereby raising the STB line and loading the phoneme. The $\overline{A}/R$ line is then driven low by the SC-01, which also lowers the STB line.

NOTE: For Complete timing details for either method of control, refer to Table 3 (Timing Specifications) and Figure 3 (Timing Diagram).

INFLECTION:

It is probably desirable during normal operation, that the two inflection inputs to the SC-01 Speech Chip both be held low. However, these inputs may be altered at any time by the user. This is accomplished by placing the desired inflection code on lines I1 and I2, and can be done in either Mode 1 or Mode 2. When a phoneme code is strobed into the SC-01, the new inflection code will also be latched. This inflection will remain in effect until the states of the I1, I2 lines are altered.

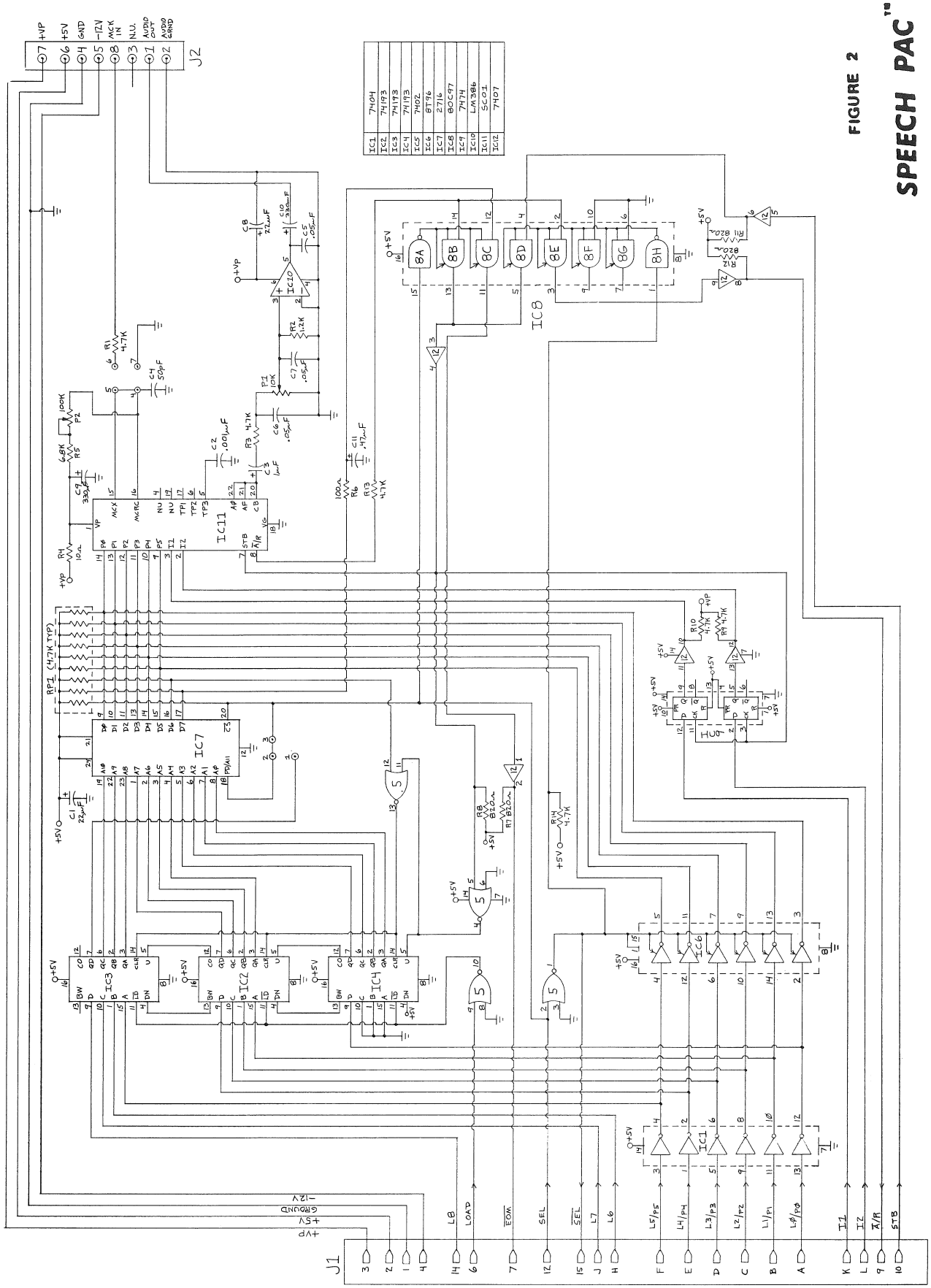
CONTROLS:

There are two external controls available on the SPEECH-PAC™. These are:

a) Volume Control: This is the knurled wheel (P1) located next to the Post Connector (see Figure 1). It is used to control the volume level of the output speech.

b) Master Clock Frequency: This is the adjustable pot (P2) located near the center of the SPEECH-PAC™ near the EPROM (see Figure 1). Altering the setting of this control alters the frequency of the Master Clock. This, in turn, alters the rate and pitch of the output speech.

NOTE: On initial power up only it might be necessary to start in Mode 2 and Strobe 'STB' line once before proceeding to Mode 1.


SPEECH PAC"
FIGURE 2
J1
J2
IC1 7404
IC2 74193
IC3 74193
IC4 74193
IC5 7402
IC6 8T96
IC7 2716
IC8 80C97
IC9 7474
IC10 LM386
IC11 SC01
IC12 7407
+VP
+5V
GND
-12V
MCK IN
N.U.
AUDIO OUT
AUDIO GND
+12V
GROUND
+5V
+VP
L8
LOAD
EOM
SEL
SEL
L7
L6
L5/P5
L4/P4
L3/P3
L2/P2
L1/P1
L0/P0
I1
I2
A/R
STB