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

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.

NC = No Connection
TPX = No Connection

Figure 3. SC-01 Footprint and Outline Dimensions
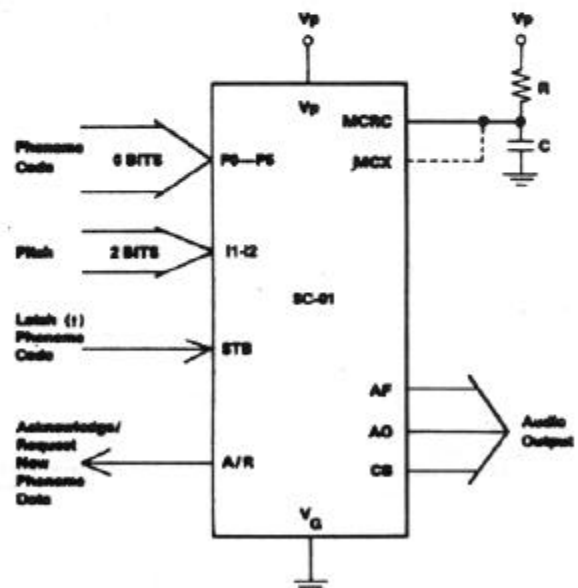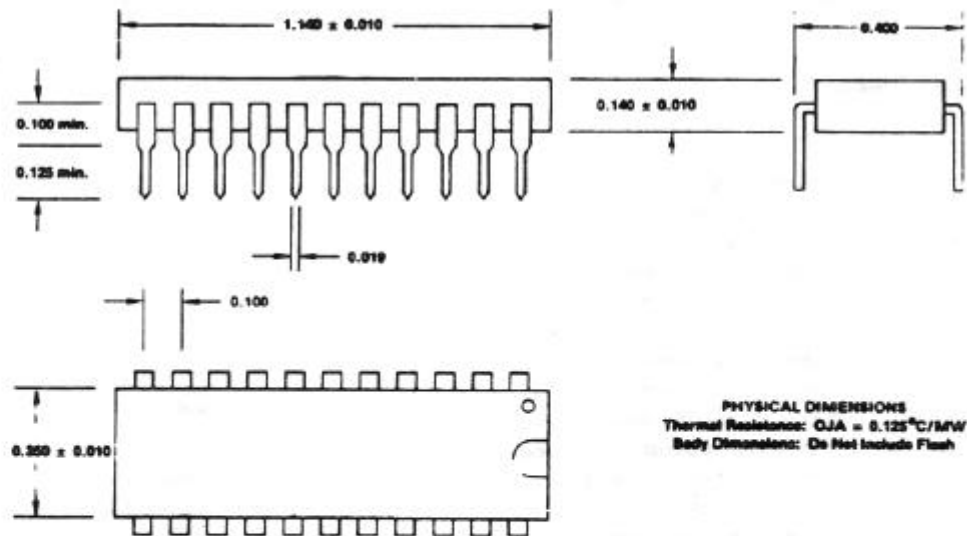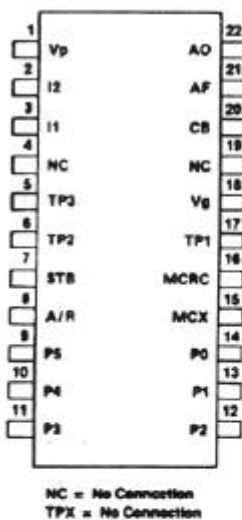
Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.

Table 1. Phoneme Chart

Phoneme Code		Phoneme Symbol		Duration (ms)	Example Word
00	0	EH3	E	59	jacket
01	1	EH2		71	enlist
02	2	EH1		121	heavy
03	3	PA0		47	no sound
04	4	DT		47	butter
05	5	A2		71	made
06	6	A1		103	made
07	7	ZH		90	azure
08	8	AH2	A	71	honest
09	9	I3		55	inhibit
0A	10	I2		80	inhibit
0B	11	I1		121	inhibit
0C	12	M	M	103	mat
0D	13	N	N	80	sun
0E	14	B	B	71	bag
0F	15	V	V	71	van
10	16	CH*		71	chip
11	17	SH		121	shop
12	18	Z	Z	71	zoo
13	19	AW1		146	lawful
14	20	NG		121	thing
15	21	AH1		146	father
16	22	OO1		103	looking
17	23	OO		185	book
18	24	L	L	103	land
19	25	K	C,K,X,Q	80	trick
1A	26	J*		47	judge
1B	27	H	H,J	71	hello
1C	28	G	G	71	get
1D	29	F	F	103	fast
1E	30	D	D	55	paid
1F	31	S	S	90	pass

Phoneme Code		Phoneme Symbol		Duration (ms)	Example Word
20	32	A		185	day
21	33	AY		65	day
22	34	Y1		80	yard
23	35	UH3		47	mission
24	36	AH		250	mop
25	37	P	P	103	past
26	38	O	O	185	cold
27	39	I		185	pin
28	40	U		185	move
29	41	Y	Y	103	any
2A	42	T	T	71	tap
2B	43	R	R	90	red
2C	44	E	I	185	meet
2D	45	W	W	80	win
2E	46	AE		185	dad
2F	47	AE1		103	after
30	48	AW2		90	salty
31	49	UH2		71	about
32	50	UH1		103	uncle
33	51	UH		185	cup
34	52	O2		80	for
35	53	O1		121	aboard
36	54	IU		59	you
37	55	U1	U	90	you
38	56	THV		80	the
39	57	TH		71	thin
3A	58	ER		146	bird
3B	59	EH		185	get
3C	60	E1		121	be
3D	61	AW		250	call
3E	62	PA1		185	no sound
3F	63	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	PA0
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						

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 sets pitch level of voiced phonemes.

Acknowledge/Request ($\overline{A}$/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 ($\overline{A}$/R Connected to STB) *	T$_{ARW}$	1	1.3	2	µs
STB Width	T$_{SW}$	200			NS
STB Low *	T$_{SL}$				NS
Propagation Delay (STB to A/R after T$_{ARW}$)	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 $\overline{A}$/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*

* *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.
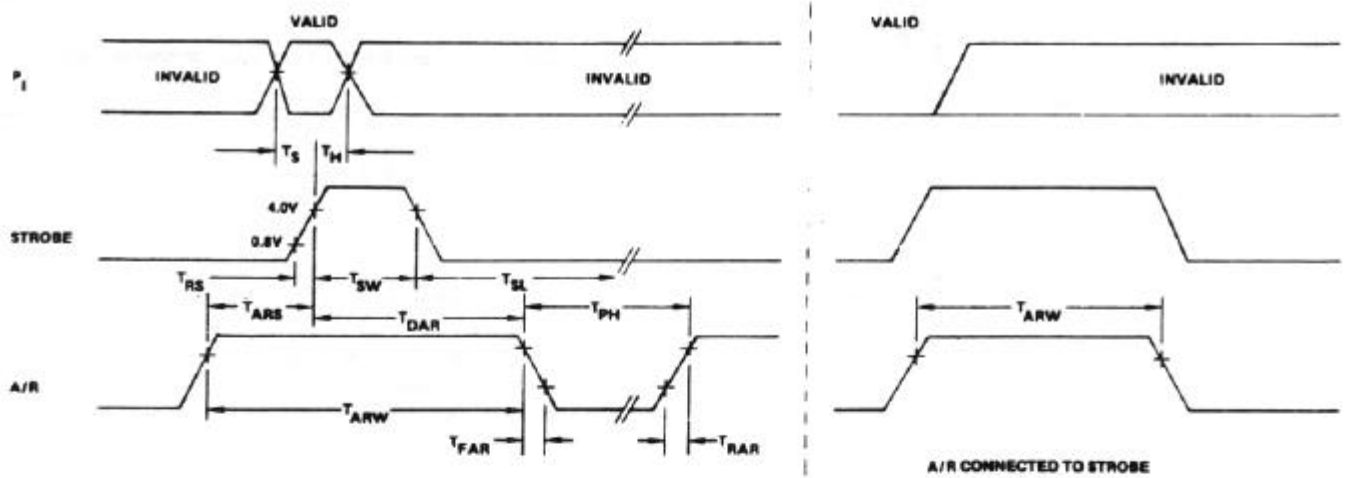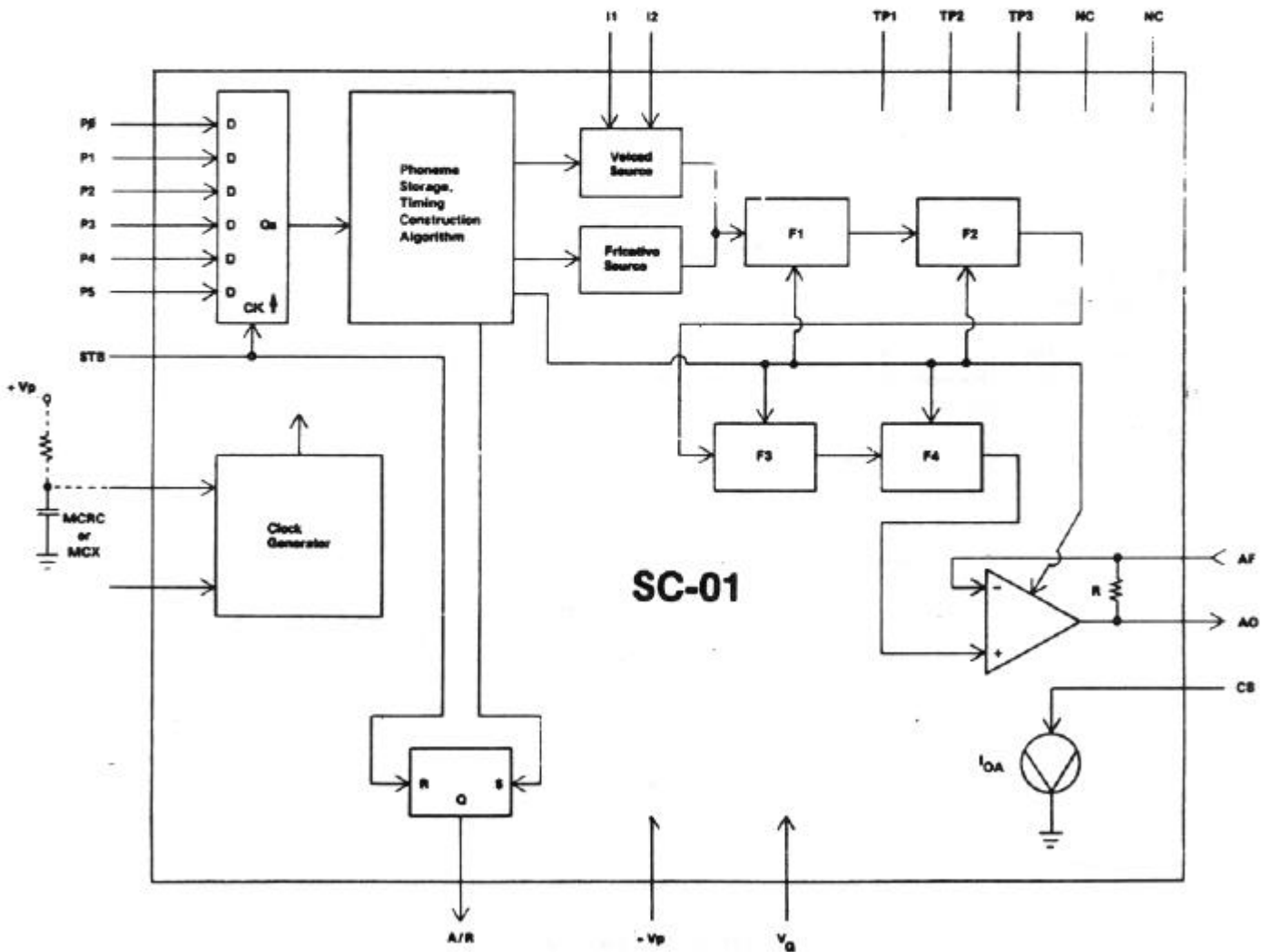
Figure 4. Timing Diagram

SC-01

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.
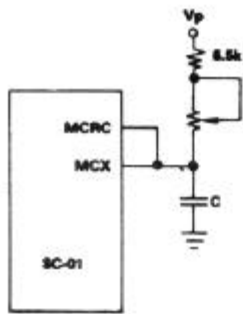
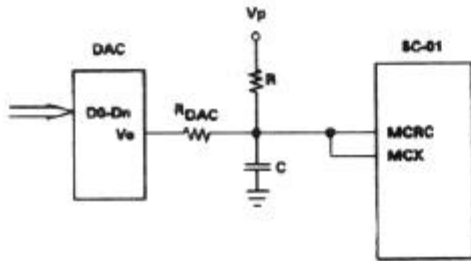Figure 6. Variable Voice by Potentiometer Control

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 u' capacitor to TP3 to insure power up reset of SC-01.

NOTE

Data at address $\emptyset$ 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

Votrax® reserves the right to alter its product line at any time, or change specifications or design without notice and without obligation.
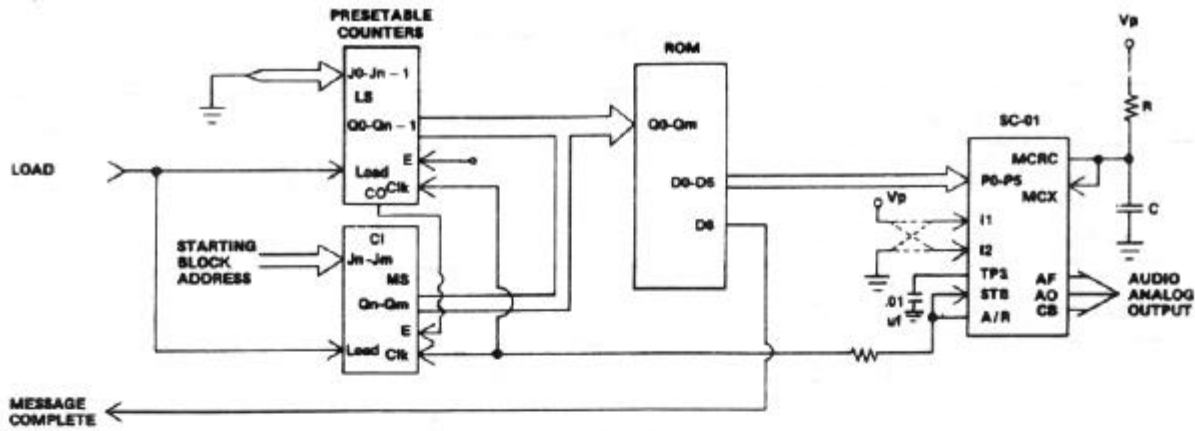
Figure 9. Multiple Message, Fixed Block Size

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 $Vp = +12.0$ 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.
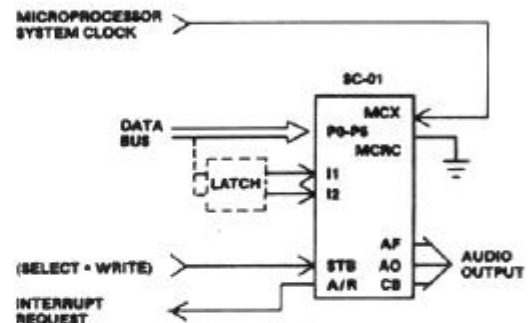

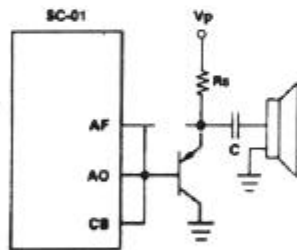Figure 10. Multiple Message, Variable Block Size
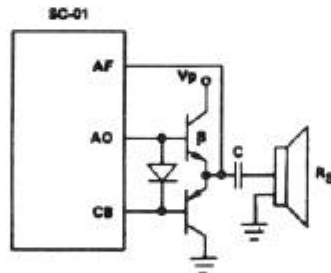

Figure 11. Class A Amplifier
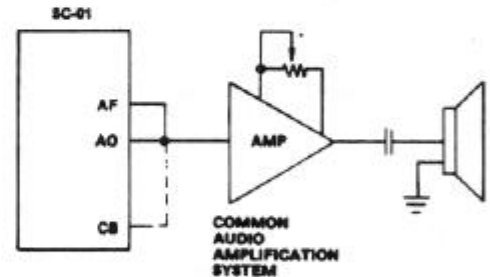

Figure 12. Class B Amplifier*

Figure 13. Controlling Audio Output Power

*For Class B Amplifier: $(\beta) \times (R_s \text{ min.}) = 81.6 \times (Vp)$ 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 × Vp	.26 × Vp	Vp-p
Output Bias Current ** (.6V < CB< Vp)	3.5	7.3	mA

ELECTRICAL CHARACTERISTICS: T_o = 0 to 70°C, Vp = 7 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 "0" (except I1, I2, MCX)		V_G - 0.5		V_G + 0.8	V_{DC}
Digital Input Logic "0" (MCX)				V_G + 1.0	V_{DC}
Digital Input Logic "0" (I1, I2)				.2 × Vp	V_{DC}
Digital Input Logic "1" (except I1, I2, MCX)		V_G + 4.0		Vp + 0.5	V_{DC}
Digital Input Logic "1" (I1, I2)		.8 × Vp			V_{DC}
Digital Input Logic "1" (MCX)		4.6			V_{DC}
Digital Output Logic "0" (I sink = 0.8mA)				V_G + 0.5	V_{DC}
Digital Output Logic "1" (I source = 0.5mA)		Vp−0.5			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)***				300	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	V_p	20	V_{DC}
Power Dissipation at 25°C	P_{DM}	650	mW
Derating Above 25°C		5	mW $^\circ$C
Operating Ambient Temperature	T_o	0 to 70	$^\circ$C
Storage Temperature	T_{STG}	−55 to 125	$^\circ$C
Input Voltage	V_{INM}	−0.5 to Vp+0.5	V_{DC}
DC Current Max. Above Vp+0.5V	I_{INM}	1.0	ma
Lead Temperature (soldering 10 sec.)	T_L	300	$^\circ$C

*_Operation above these limits could damage the device._

NORMAL OPERATING CONDITIONS: $7v \leqslant V_p \leqslant 14v, 0^\circ C \leqslant T_o \leqslant 70^\circ C$