

DIGITALKER™ Speech System

BLX-281	Speech Synthesis Expansion Module
DT1000	DIGITALKER™ Speech Synthesis Evaluation Board
DT1050/DT1053	DIGITALKER™ Standard Vocabulary Kit
DT1051/DT1054	DIGITALKER™ Speech Evaluation Kit
DT1052/DT1055	DIGITALKER™ Basic Numbers Kit
DT1056/DT1057	DIGITALKER™ Standard Vocabulary Kit
MM54104	DIGITALKER™ Speech Synthesis System
LB-54	Circuit for Evaluation of Custom Vocabulary EPROM
	Prototype Set
AN-252	Speech Synthesis



Section 13
DIGITALKER™
Speech Synthesis

13



Section Contents

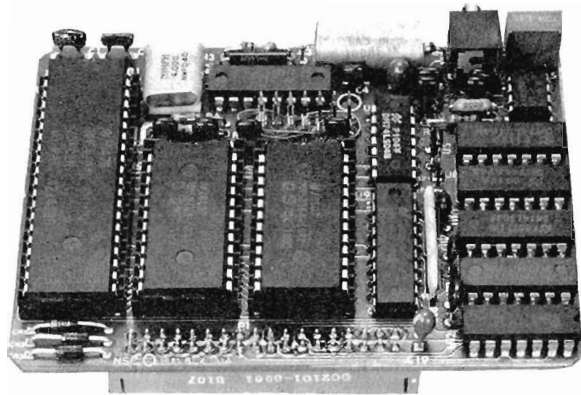
BLX-281 Speech Synthesis Expansion Module	13-3
DT1000 DIGITAL TALKER™ Speech Synthesis Evaluation Board	13-7
DT1050/DT1053 DIGITAL TALKER™ Standard Vocabulary Kit	13-14
DT1051/DT1054 DIGITAL TALKER™ Speech Evaluation Kit	13-22
DT1052/DT1055 DIGITAL TALKER™ Basic Numbers Kit	13-24
DT1056/DT1057 DIGITAL TALKER™ Standard Vocabulary Kit	13-26
MM54104 DIGITAL TALKER™ Speech Synthesis System	13-34
LB-54 Circuit for Evaluation of Custom Vocabulary EPROM Prototype Set	13-41
AN-252 Speech Synthesis	13-43

National Semiconductor

DIGITALKER™ Speech Synthesis

BLX-281

Speech Synthesis Expansion Module



- BLX bus-compatible I/O expansion
- Speech synthesis based on DIGITALKER™
- Large vocabulary adequate for most applications
- On-board filter and half-watt amplifier
- Simple operation for user
 - I/O write with word/sound address
 - Interrupt asserted when complete
- BLX bus on-board expansion eliminates Multibus™ system bus latency and increases system throughput

Product Overview

The BLX-281 Speech Synthesis Expansion Module is a member of the new line of BLX bus-compatible expansion module products from National Semiconductor Corporation. The BLX-281 plugs directly into any BLX bus-compatible host board offering low cost incremental on-board expansion. As a result, any BLX bus-compatible host board may be given the ability to "speak". By merely adding a speaker to a system containing the BLX-281, many users can do away with CRTs, printers, rows of LEDs, or similar communications devices. This lowers the cost of most systems, and has the added benefit of removing messages which are potentially ambiguous and hard-to-understand for untrained users. The BLX-281 contains 144 words, sounds, tones, and durations of silence, each of which has a unique address. A table of addresses (desired words/sounds) is built, and passed to the BLX-281. An on-board filter and amplifier provide the actual speech signal to a standard miniature phone jack. The BLX module is closely coupled to the host board through the BLX bus, and as such, offers maximum on-board performance, and frees Multibus system traffic for other system resources. Incremental power dissipation is minimal, requiring only 3.7 watts.

DIGITALKER is a trademark of National Semiconductor Corp.

Functional Description

The BLX-281 Speech Synthesis Expansion Module uses the MM54104 Speech Processor Chip from National Semiconductor Corporation. The digitized and compressed speech data are contained in an MM52164 Maxi-ROM. The system software communicates with the BLX-281 across the BLX bus using I/O read/write commands.

Vocabulary

The standard vocabulary set offered on the BLX-281 is shown in Table I, along with the assigned addresses for each item. By combining the appropriate words, sounds, tones, and silence durations, speech can be generated to satisfy many applications.

Words required, but not found in the table, can frequently be built. Examples of this are: combine "RE" with "SET" for "RESET", or combine "DEGREE" with "SS" for "DEGREES".

In normal human speech, the brain puts durations of silence between the words to make the sentence flow smoothly. This is provided for in the BLX-281 (see Table I). A suggestion for improved phrase quality is to insert 80 milliseconds of silence prior

Multibus is a trademark of Intel Corp.

to words beginning with the letters K, T, P, B, D, and G, and to add 40 milliseconds of silence after words ending in those same letters.

The "voice" output of the BLX-281 is a highly intelligible, male voice. If another voice is required, or the application is non-English, or involves unusual terminology, any voice can be processed for use on the BLX-281 by the factory.

Host Interface

The BLX bus-compatible host board merely passes the address of the desired word/sound to the BLX-281 Speech Synthesis Expansion Module via an I/O write. When the operation is complete, an Interrupt is generated. This informs the host of the end of the speech

sequence, and allows for cascading of addresses for true, human-quality sentences.

Interrupt Requests

There is one interrupt line from the Speech Processor Chip that generates an interrupt request to the host CPU. It is active on completion of each speech sequence. It is cleared by an I/O read to the BLX-281.

Installation

The BLX-281 module plugs directly into either of the female BLX connectors on the host board. The module is then secured at one additional point with nylon hardware to insure the mechanical security of the assembly (see Figures 1 and 2).

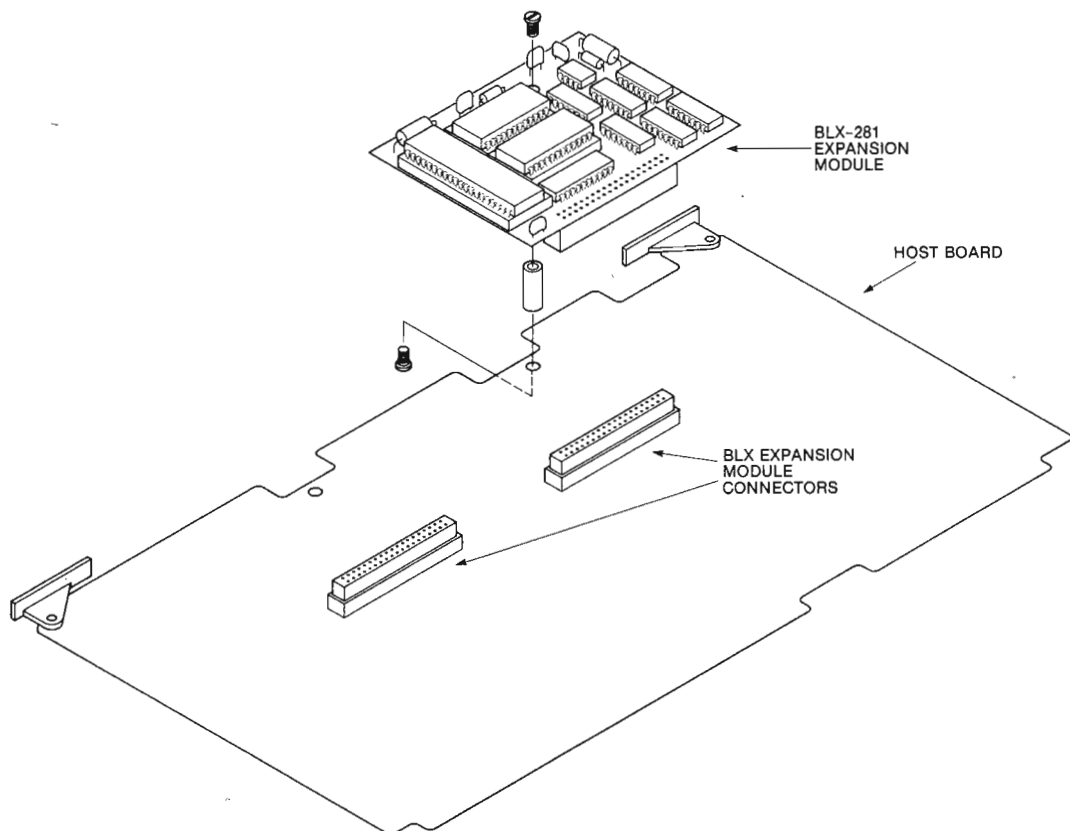


Figure 1. Installation of the BLX-281 Module on a Host Board

Table I. Master Word List

Word	8-Bit Binary Address		Word	8-Bit Binary Address		Word	8-Bit Binary Address	
	Bit 7	Bit 0		Bit 7	Bit 0		Bit 7	Bit 0
THIS IS DIGITALKER	0	0	Q	0	0	IS	0	1
ONE	0	0	R	0	0	IT	0	1
TWO	0	0	S	0	0	KILO	0	1
THREE	0	0	T	0	0	LEFT	0	1
FOUR	0	0	U	0	0	LESS	0	1
FIVE	0	0	V	0	0	LESSER	0	1
SIX	0	0	W	0	0	LIMIT	0	1
SEVEN	0	0	X	0	0	LOW	0	1
EIGHT	0	0	Y	0	0	LOWER	0	1
NINE	0	0	Z	0	0	MARK	0	1
TEN	0	0	AGAIN	0	0	METER	0	1
ELEVEN	0	0	AMPERE	0	0	MILE	0	1
TWELVE	0	0	AND	0	0	MILLI	0	1
THIRTEEN	0	0	AT	0	0	MINUS	0	1
FOURTEEN	0	0	CANCEL	0	0	MINUTE	0	1
FIFTEEN	0	0	CASE	0	0	NEAR	0	1
SIXTEEN	0	0	CENT	0	1	NUMBER	0	1
SEVENTEEN	0	0	400Hz TONE	0	1	OF	0	1
EIGHTEEN	0	0	80Hz TONE	0	1	OFF	0	1
NINETEEN	0	0	20ms SILENCE	0	1	ON	0	1
TWENTY	0	0	40ms SILENCE	0	1	OUT	0	1
THIRTY	0	0	80ms SILENCE	0	1	OVER	0	1
FORTY	0	0	160ms SILENCE	0	1	PARENTHESIS	0	1
FIFTY	0	0	320ms SILENCE	0	1	PERCENT	0	1
SIXTY	0	0	CENTI	0	1	PLEASE	0	1
SEVENTY	0	0	CHECK	0	1	PLUS	0	1
EIGHTY	0	0	COMMA	0	1	POINT	0	1
NINETY	0	0	CONTROL	0	1	POUND	0	1
HUNDRED	0	0	DANGER	0	1	PULSES	0	1
THOUSAND	0	0	DEGREE	0	1	RATE	0	1
MILLION	0	0	DOLLAR	0	1	RE	0	1
ZERO	0	0	DOWN	0	1	READY	0	1
A	0	0	EQUAL	0	1	RIGHT	1	0
B	0	0	ERROR	0	1	SS	1	0
C	0	0	FEET	0	1	SECOND	1	0
D	0	0	FLOW	0	1	SET	1	0
E	0	0	FUEL	0	1	SPACE	1	0
F	0	0	GALLON	0	1	SPEED	1	0
G	0	0	GO	0	1	STAR	1	0
H	0	0	GRAM	0	1	START	1	0
I	0	0	GREAT	0	1	STOP	1	0
J	0	0	GREATER	0	1	THAN	1	0
K	0	0	HAVE	0	1	THE	1	0
L	0	0	HIGH	0	1	TIME	1	0
M	0	0	HIGHER	0	1	TRY	1	0
N	0	0	HOUR	0	1	UP	1	0
O	0	0	IN	0	1	VOLT	1	0
P	0	0	INCHES	0	1	WEIGHT	1	0

Note: Address 8F_H is the last legal address in this word list. Exceeding address 8F_H will produce pieces of unintelligible, invalid speech data.

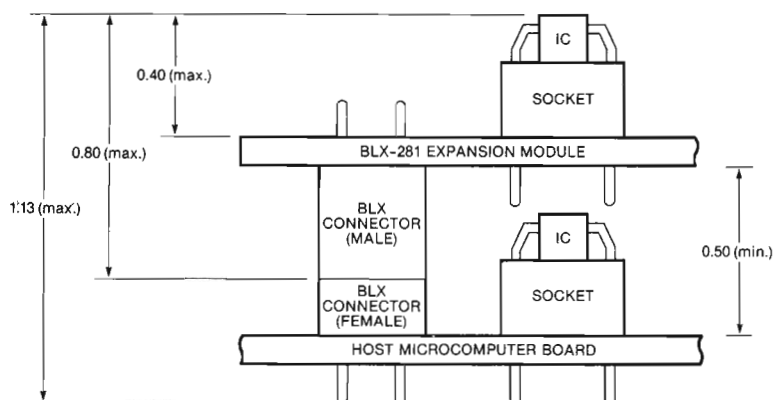


Figure 2. BLX-281 Expansion Module Mounting Clearances (inches)

Specifications

Word Size

Data — 8 bits

I/O Addressing

Function	Type of Operation	BLX Connector Port Address
Data Transfer	Write	X0-XF
Interrupt Clear	Read	X0-XF

Note: The port addresses are determined on the host BLC microcomputer. Refer to the Hardware Reference Manual for your host BLC microcomputer to determine the first digit (X) of the connector port address.

Vocabulary — See Table I

Interrupts — One interrupt request at end of speech sequence

Interfaces — BLX Bus — All signals TTL compatible
Speaker Port — $\frac{1}{2}W$ audio signal into 4-8 Ω

Speaker Port Connector — Standard miniature phone-jack

Physical

Height: 2.85 in. (7.24 cm)

Width: 3.70 in. (9.40 cm)

Depth

BLX-281 Module

0.80 in. (2.04 cm)

BLX-281 Module + Host Board

1.13 in. (2.86 cm)

Weight: 1.7 oz. (48 gm)

Electrical

+5 VDC $\pm 5\%$ @ 385 mA

+12 VDC $\pm 5\%$ @ 150 mA

Environmental

Operating Temperature: 0°C to 55°C

Relative Humidity: 0% to 90%, non-condensing

Ordering Information

BLX-281 Speech Synthesis Expansion Module

Documentation

420306414-001 BLX-281 Speech Synthesis Expansion Module User's Manual

DT1000 DIGITALKER™ Speech Synthesis Evaluation Board

General Description

The DIGITALKER™ (DT1000) speech synthesis evaluation board is an extremely easy to use device for understanding the operation and application of the DIGITALKER chip set in an end product.

The DT1000 contains all components required to output speech upon demand: a speech processor chip (SPC), 2 MAXI-ROMs® containing 138 individual words, linear filter, audio amplifier, keyboard, and a COPS™ microcontroller complete with stored data programmed to provide the various functions on the board. The only external hardware required for complete operation are a single 7V-11V power supply, a speaker of your choice for size and quality, and this instruction sheet.

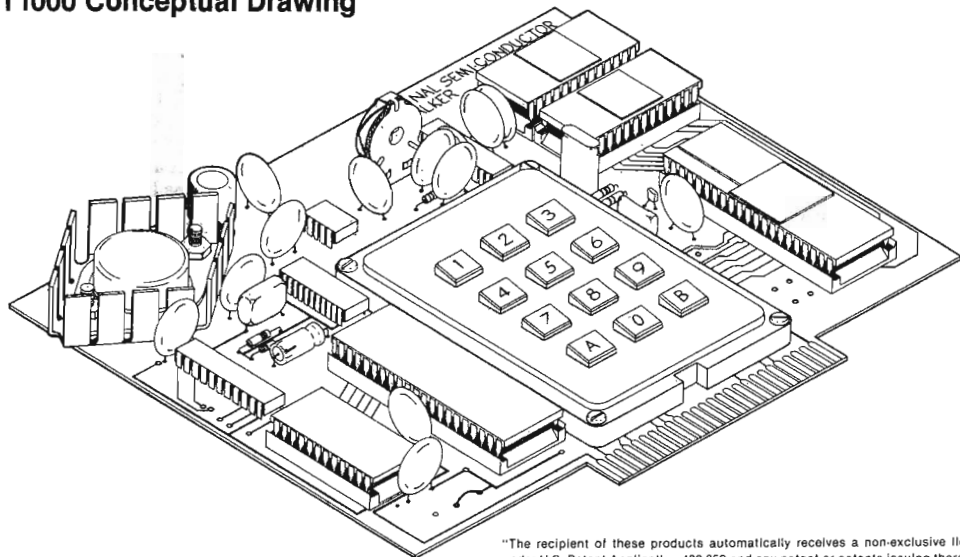
The 2 speech MAXI-ROMs employed on the board contain a brief introductory phrase, 138 separate and individual words consisting of numbers and letters of the alphabet, assorted useful nouns, verbs and tones; and 5 different individual silence durations. (In constructing a phrase, different silence durations between different words significantly affect the overall quality of the phrase.)

A COPS program is provided which permits the user to: 1) sequentially output each word automatically; 2) repeat any desired word; 3) build and store several short phrases for outputting when desired; 4) output a "canned" phrase which permits insertions and changing of a word in the phrase; 5) play a simple game which requires some interaction between the keyboard and the user; and 6) output hex equivalent decimal number inputs.

Features

- Only a single 7V-11V power supply and inexpensive loudspeaker required for total operation
- 138 individually addressable words, applicable to many products
- Programmed COPS processor permits 6 individual program modes
- Demonstrates the extreme flexibility and ease of application of the DIGITALKER chip set
- 1/2 watt audio amplifier on-board
- Edge connector facilitates tying in to external processor system (3M Company connector part number #3415)

DT1000 Conceptual Drawing



MAXI-ROM® is a registered trademark of National Semiconductor Corp.
DIGITALKER™ and COPS™ are trademarks of National Semiconductor Corp.

"The recipient of these products automatically receives a non-exclusive license under U.S. Patent Application 432,859 and any patent or patents issuing thereon to use such products, to assemble or otherwise incorporate them into further products which may be covered by said patent application, or any patent or patents issuing thereon, and to use, sell, or otherwise dispose of such products".
Protected by U.S. Pat. No. 4124125, F.M. Mozer licenses available.

Absolute Maximum Ratings

Operating Temperature Range	0°C to 70°C
V _{DD} Supply Voltage	12V

Electrical Characteristics T_A = 25°C

Symbol	Parameter	Conditions	Min	Typ	Max	Units
V _{DD} *	Supply Voltage	V _{SS} = 0V	7.0		11.0	V
I _{DD}	Supply Current	V _{DD} = 11V			0.6	A
	Amplifier Output	V _{DD} = 11V, 8Ω	0.55			W
		V _{DD} = 9V, 8Ω	0.50			W

* Important! V_{DD} must be regulated!

All pin connections, except for speaker out, and power supply, are TTL compatible.

Functional Description

INSTRUCTIONS FOR USE

In any case, plus 7V to 11V direct current and ground must be brought to the respective pins on the edge connector of the DT1000 board. See Figure 1.

BASIC MODE OPERATIONS

1. **Power Up Mode.** At power up, the DT1000 will say "This is DIGITALKER". You can make it repeat this phrase by depressing [A]. Further depressions of [A] will repeat the same phrase until a new word address is entered. From the DT1000 master word list (Table I) select the 3-digit address of the next word desired. The new word can be outputted simply by keying in its 3-digit address and ending with an [A] key depression. For example, the word "pound" will be spoken as a result of key sequence [1], [2], [3], [A]. Additional depressions of A will repeat the word "pound". To output a new word, repeat the above sequence.

If a mistake is made in the address entry of the desired word, simply continue to key in the correct address. The DT1000 only remembers the last 3 address digits keyed prior to depression of the [A] key. For example, if the desired word is "pound" and keys [1], [2], [2] were depressed, simply rekey the proper address followed by the [A] key as shown [1], [2], [3], [A].

In this Power Up Mode the [B] key will advance to the next sequential word in the DT1000 master word list, one word per depression, starting with the next sequential word after the word selected by the entry sequence shown above. Continued depressions of [B] will manually sequence through the complete word list, wrapping around from address 143 to address 000 and ending up at the first word in the sequence. Depressions of the [A] key will take the DT1000 back to the starting word and address inputted via the entry sequence shown above. Additional depressions of the [B] key would then sequence from that address again. Should a different start word be desired, you must enter the respective address of that word as is shown in the above sequence.

The DT1000 has 144 legal address locations, these are shown in the Master Word List as being addresses 000 through 143. If an address of 144 through 199 is inadvertently keyed in, a response of "Please Try Again" will be outputted. Addresses 200 and up will put the DT1000 into various operating modes as explained in paragraphs 2 through 7.

2. **Auto Repeat/Auto Cycle Mode.** The Auto Repeat/Cycle Mode is entered by depressing [2], [0], [0] and either the [A] or [B] key. An additional depression of [A] will start an automatic repeating cycle of some word and address, until the [A] key is depressed and held momentarily. Depression of the [B] key will start the automatic sequential cycling through the complete word list, beginning with the last word that was outputted in either the Auto Repeat or Auto Cycle Mode. The sequential cycling will automatically continue through the entire word list, wrapping around from address 143 to address 000 each time until the [A] or [B] key is depressed and held momentarily. When in this mode you can choose a new starting word by entering its respective address and keying either the [A] or [B] keys.

3. **Decimal to Hex Conversion Mode.** Key sequence [3], [0], [0], [A] (or [B]) will put the DT1000 into a mode where it will automatically convert a decimal address into its hexadecimal equivalent. Once in this mode, key in the decimal address desired from the master word list, and depress the [A] key twice. The respective word will be spoken on the first depression of [A], a further depression of [A] will cause its hexadecimal equivalent address to be spoken.

A [B] key activation stand alone will set the DT1000 into the Power Up Manual Control Mode.

Functional Description (Continued)

- 4. Phrase Construct Mode.** In this mode, it is possible to string respective word addresses together to create up to 3 phrases or short sentences and play them back upon demand. The phrase modes are **[4][0][0]**, **[5][0][0]**, and **[6][0][0]**. The 400 mode will store up to twelve word addresses, the 500 and 600 modes will store up to six addresses each. Initial phrase construction should begin with key sequence **[4][0][0]** where the **[0]** depression clears out any addresses previously stored. (In either the **[4][0][0]**, **[5][0][0]**, **[6][0][0]** sequence, the B key will "clear" out any previous addresses stored in those respective modes.) You are then ready to string addresses together to construct a phrase. To construct the phrase, simply enter each address of the words desired in the phrase and in the sequence in which the words are to be outputted. Each word address keyed in must be followed by two depressions of the **[0]** key. The first depression of the **[0]** key will speak the word at that respective address to give you a chance to "hear" the word before its address is "loaded" into the DT1000 RAM. A second depression of the **[0]** key will store that address into the DT1000 RAM and that word is then in position in the desired phrase. If upon the first depression of the **[0]** key the word outputted is not the desired word, simply key in the correct address and depress the **[0]** key again, and again finally to store into the RAM. In the construction of a phrase a "beep" will occur during the addressing of the last possible address that will fit into that particular phrase mode. (If additional addresses are still keyed in, they will replace the first addresses loaded in that same phrase.)

An example of constructing a phrase is as follows for the desired word sequence "To start the time set the meter and go", key sequence: **[4][0][0][0][0][2][0][0][1][3][5][0][0][1][3][8][0][0][1][3][9][0][0][1][3][1][0][0][1][3][8][0][0][1][0][6][0][0][0][6][0][0][0][8][6][0][0]**. To output this same phrase, depress the **[A]** key. The above phrase should automatically be spoken from the DT1000. Additional keyings of **[A]** will output the same phrase until it is cleared out by a **[4][0][0][0]** sequence.

Registers 500 and 600 can be loaded in exactly the same way as shown above, except that the **[5][0][0]** and/or the **[6][0][0]** keys must be addressed to load those respective registers. Remember that registers 500 and 600 are each only 6 addresses long.

If you chose to construct the exact phrase as shown above, you may have noticed that the spoken output was rather a mechanical output. This is primarily due to the fact that the words were butted against each other. In normal human speech, the brain puts durations of silence between the words to make the sentence flow smoothly. Since several durations of silence are provided in the Master Word List, the actual quality of the same phrase can be significantly improved by adding durations of silence (also assigned addresses) between the words. As one thinks about how the phrase is actually spoken, one might assume the approximate duration of silence between each word, and insert the closest duration of silence from the word list. It is found that some experimentation in this area can significantly enhance the quality of the desired phrase. A hint in

this area would be that for words beginning with the letters K, T, P, B, D, and G insert 80 milliseconds silence prior to the words, and for words ending in the same letters as above, 40 milliseconds silence following the word is recommended. It is also possible in this mode to make any singular word plural by the addition of "SS" (Address 129) to the word. In this case no silence should be inserted between the word and the "SS".

- 4A. Phrase Output Mode.** As stated in (4) any phrase can be outputted by being in the 400, 500 or 600 modes and depressing the **[A]** key. It is also possible to output all 3 phrases in any sequence. To "string" these phrases together, simply key in the phrase sequence desired concluded with two depressions of the **[A]** key. Key sequence **[4][5][6][A][A]** would output phrases at 400-500 and 600 respectively. (This would indicate that a sentence 24 addresses long might be constructed.) Any phrase sequence might actually be chosen, 546, 645, etc. For an interesting effect the same phrase could be outputted twice or even three times such as 455, 444, 664, etc.

- 5. Canned Phrase Mode.** Key sequence **[7][0][0][A]** (or **[B]**) will output a fixed phrase "The time is ____ P.M.". This gives you the ability to insert the desired word(s) in the blank location. In this case "twelve OH one" might be appropriate. While in the **[7][0][0][A]** mode simply key in the respective addresses of the words desired, inserting silences if required, exactly the same as constructing a phrase in the 400-500 or 600 modes. To output the completed phrase, simply depress the **[A]** key. To insert a new word sequence into the blank, key **[7][0][0][B]** to erase the original contents. Then enter the new word addresses as required.

This mode demonstrates how a talking clock or a trip computer might work. Changing data can be inserted at the required time as a part of a fixed message.

- 6. Reaction Timer Game.** Key sequence **[8][0][0][A]** (or **[B]**) enters you into a simple game which could conceptually be a real product. In this mode, the DT1000 speaks ten random numbers from zero to nine, with a pause between each number output. The game is to hit the respective key as fast as possible after the number has been called out. After the tenth and last number has been depressed, a tone is outputted and the total reaction time is spoken as "seven point three two five seconds". Obviously, the game is to have the lowest possible total reaction time. *Note that it is necessary to eventually hit the correct key for the number called out.* If the wrong key is depressed, the DT1000 will not output another number until the correct key is depressed—meanwhile, time is accruing. It should also be noted that the random pauses between word outputs is not part of total elapsed time.

To continue playing the game, keying the **[A]** key will output a new set of numbers. To exit the game mode, depress the **[B]** key.

- 7. Back to Power Up Mode.** Key sequence **[9][0][0][A]** (or **[B]**) will put the DT1000 back into the Power Up Mode. Refer to mode (1.) explained in the earlier section of this data sheet for all operations covered by this mode.

Functional Description (Continued)

GENERAL COMMENTS

1. The DT1000 is always in one of the modes. To exit a mode, simply key in the control code of the next desired mode.
2. "SS" (located at address 129) can make singular words plural.
3. "Centi", "milli", "re" are prefixes to make words like "centimeter", "reset", etc. Some prefixes do not blend well directly with some words such as "milli ampere". In these cases, insert an appropriate amount of silence between the words.
4. High output volume can be obtained by supplying 11V to power supply input (pin 50).

SPEECH QUALITY

The actual speech quality of the DT1000 is affected by many factors. Certainly the quality is affected most significantly by the actual speaker and baffle chosen to output the final speech data. Although the DT1000 will drive most any size of "common" PM speakers, care ought to be made in the actual selection of the speaker, AND its respective baffle or enclosure. An unbaffled speaker will not give a true response, small speakers typically do not reproduce low frequencies. Truly, the most desirable com-

bination for best quality would be a medium size speaker 6 inches to 12 inches in diameter, and housed in a solid wood enclosed baffle.

One can actually "experiment" with the quality by trying various speaker and baffle combinations.

APPLICATION WITH EXTERNAL PROCESSOR

The DT1000 is designed so that it is possible to access only the DIGITALKER portion of the board. *Important: it is necessary to remove the COP402 from the DT1000 in this mode.* The DIGITALKER portion is defined as the speech processor chip (SPC) and the speech ROM(s) which contain the actual vocabulary (see Table I). The inputs required to connect the DIGITALKER (and the vocabulary of the DT1000 board) to an external processor have been made available on the pin edge connector (refer to Figure 3).

The following describes the function of all speech processor chip (SPC) inputs and outputs, and all other inputs and outputs required for operation in this external processor mode. Note: in the following descriptions and Table I, a low represents a logic zero (0.4V nominal) and a high represents a logic one (2.4V nominal).

CONNECTION REQUIREMENTS FOR EXTERNAL PROCESSOR APPLICATIONS

Edge Connector Pin Number	Function	Edge Connector Pin Number	Function
8	Chip Select ($\overline{CS}$). The SPC is selected when $\overline{CS}$ is low. It is only necessary to have $\overline{CS}$ low during a command to the SPC. It is not necessary to hold $\overline{CS}$ low for the duration of the speech data.	37	Write Strobe ($\overline{WR}$). This line latches the starting address (A0-A7) into a register. On the rising edge of the $\overline{WR}$, the SPC starts execution of the command specified by CMS. The command sequence is shown in the timing waveform section. If a command to start a new speech sequence is issued during a speech sequence, the new speech sequence will be started immediately.
3	Data Bus (SW1-SW8). SW1 (LSB) This is an 8-bit parallel binary data bus which accepts the binary address of the desired word. The binary addresses are available from Table I and are the same as the decimal address from the word list. Unused inputs must be connected to ground when used with external logic.	50	Power Supply Voltage (V_{DD}). Plus 7V to 11V maximum, direct current, to SPC, filter and amplifier sections. Important! V_{DD} must be regulated!
43		47	Ground (V_{SS}).
45		7	Interrupt Output (INTR). This signal goes high at the completion of any speech sequence. It is reset by the next valid command. It is also reset at power up.
40		10	Speaker Output. 4 Ω -8 Ω 1/2W at V_{DD} , 11.0V
13			
9			
16			
11			
34	Command Select (CMS). This line is used to specify the two commands to the SPC.		
	CMS Command		Function
	0		Reset interrupt and start speech sequence
	1		Reset interrupt only

Functional Description (Continued)

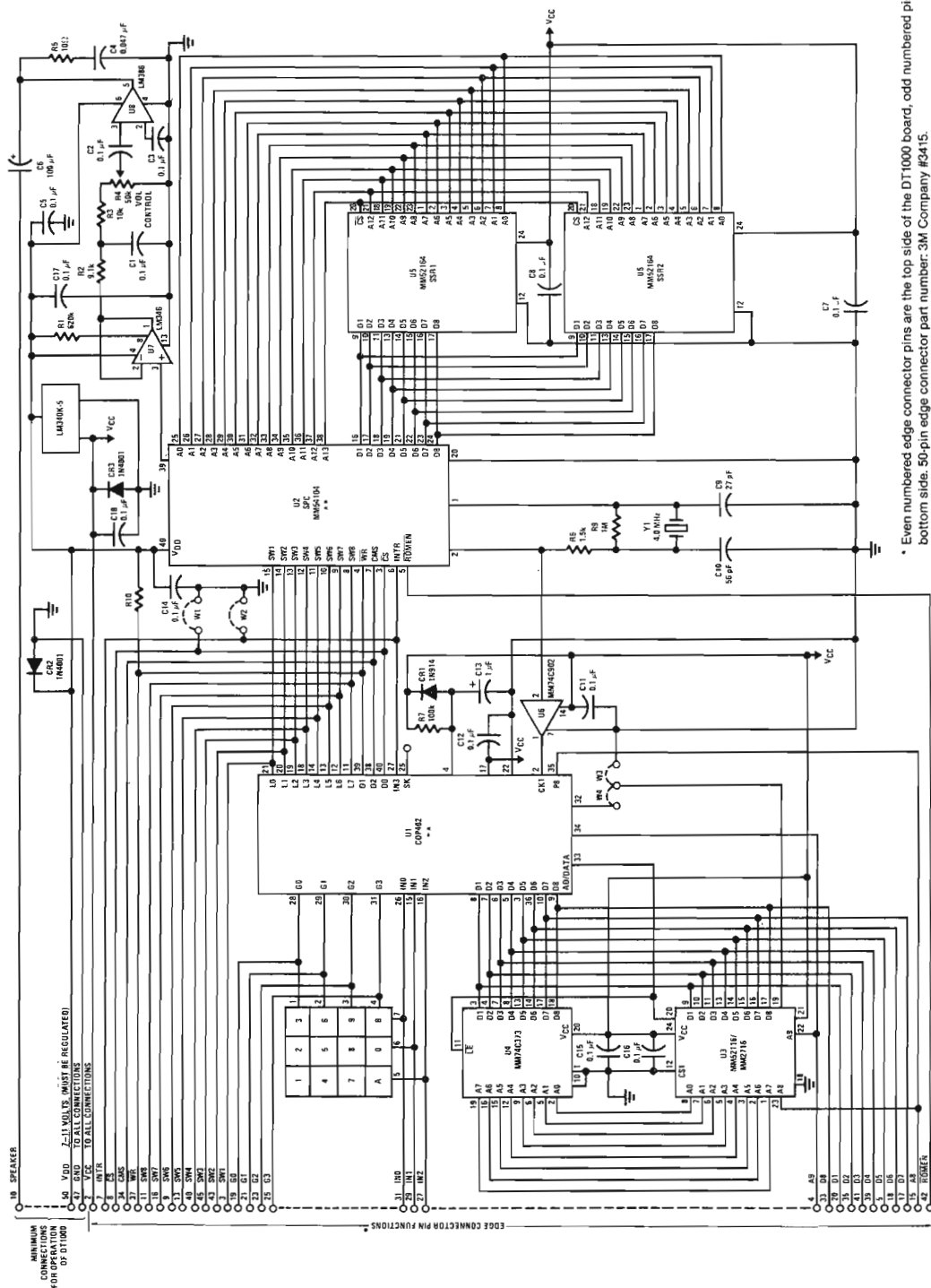
TABLE I. DT1000 MASTER WORD LIST

Word	Keyboard Address	8-Bit Binary Address		Word	Keyboard Address	8-Bit Binary Address		Word	Keyboard Address	8-Bit Binary Address	
		SW 8	SW 1			SW 8	SW 1			SW 8	SW 1
THIS IS DIGITALKER	000	0	0	W	054	0	1	MILLI	108	0	1
ONE	001	0	0	X	055	0	1	MINUS	109	0	1
TWO	002	0	0	Y	056	0	1	MINUTE	110	0	1
THREE	003	0	0	Z	057	0	1	NEAR	111	0	1
FOUR	004	0	0	AGAIN	058	0	1	NUMBER	112	0	1
FIVE	005	0	0	AMPERE	059	0	1	OF	113	0	1
SIX	006	0	0	AND	060	0	1	OFF	114	0	1
SEVEN	007	0	0	AT	061	0	1	ON	115	0	1
EIGHT	008	0	0	CANCEL	062	0	1	OUT	116	0	1
NINE	009	0	0	CASE	063	0	1	OVER	117	0	1
TEN	010	0	0	CENT	064	0	1	PARENTHESIS	118	0	1
ELEVEN	011	0	0	400HERTZ TONE	065	0	1	PERCENT	119	0	1
TWELVE	012	0	0	80HERTZ TONE	066	0	1	PLEASE	120	0	1
THIRTEEN	013	0	0	20MS SILENCE	067	0	1	PLUS	121	0	1
FOURTEEN	014	0	0	40MS SILENCE	068	0	1	POINT	122	0	1
FIFTEEN	015	0	0	80MS SILENCE	069	0	1	POUND	123	0	1
SIXTEEN	016	0	0	160MS SILENCE	070	0	1	PULSES	124	0	1
SEVENTEEN	017	0	0	320MS SILENCE	071	0	1	RATE	125	0	1
EIGHTEEN	018	0	0	CENTI	072	0	1	RE	126	0	1
NINETEEN	019	0	0	CHECK	073	0	1	READY	127	0	1
TWENTY	020	0	0	COMMA	074	0	1	RIGHT	128	1	0
THIRTY	021	0	0	CONTROL	075	0	1	SS (Note 1)	129	1	0
FORTY	022	0	0	DANGER	076	0	1	SECOND	130	1	0
FIFTY	023	0	0	DEGREE	077	0	1	SET	131	1	0
SIXTY	024	0	0	DOLLAR	078	0	1	SPACE	132	1	0
SEVENTY	025	0	0	DOWN	079	0	1	SPEED	133	1	0
EIGHTY	026	0	0	EQUAL	080	0	1	STAR	134	1	0
NINETY	027	0	0	ERROR	081	0	1	START	135	1	0
HUNDRED	028	0	0	FEET	082	0	1	STOP	136	1	0
THOUSAND	029	0	0	FLOW	083	0	1	THAN	137	1	0
MILLION	030	0	0	FUEL	084	0	1	THE	138	1	0
ZERO	031	0	0	GALLON	085	0	1	TIME	139	1	0
A	032	0	1	GO	086	0	1	TRY	140	1	0
B	033	0	1	GRAM	087	0	1	UP	141	1	0
C	034	0	1	GREAT	088	0	1	VOLT	142	1	0
D	035	0	1	GREATER	089	0	1	WEIGHT (Note 2)	143	1	0
E	036	0	1	HAVE	090	0	1				
F	037	0	1	HIGH	091	0	1				
G	038	0	1	HIGHER	092	0	1				
H	039	0	1	HOURLY	093	0	1				
I	040	0	1	IN	094	0	1				
J	041	0	1	INCHES	095	0	1				
K	042	0	1	IS	096	0	1				
L	043	0	1	IT	097	0	1				
M	044	0	1	KILO	098	0	1				
N	045	0	1	LEFT	099	0	1				
O	046	0	1	LESS	100	0	1				
P	047	0	1	LESSER	101	0	1				
Q	048	0	1	LIMIT	102	0	1				
R	049	0	1	LOW	103	0	1				
S	050	0	1	LOWER	104	0	1				
T	051	0	1	MARK	105	0	1				
U	052	0	1	METER	106	0	1				
V	053	0	1	MILE	107	0	1				

Note 1: "SS" makes any singular word plural.

Note 2: Address 143 is the last legal address in this particular word list. Exceeding address 143 in an external processor application will produce pieces of unintelligible invalid speech data.

DT1000 Schematic Diagram



- Even numbered edge connector pins are the top side of the DT1000 board, odd numbered pins are on bottom side. 50-pin edge connector part number: 3M Company #3415.

- The COP402 device must be removed from the socket for external processor applications.

AC Electrical Characteristics

(For Use in External Processor Application) $T_A = 0^\circ\text{C}$ to 70°C , $V_{DD} = 7\text{V}$ – 11V , $V_{SS} = 0\text{V}$, unless otherwise specified.

Symbol	Parameter	Min	Max	Units
t_{aw}	CMS Valid to Write Strobe	350		ns
t_{csw}	Chip Select ON to Write Strobe	310		ns
t_{dw}	Data Bus Valid to Write Strobe	50		ns
t_{wa}	CMS Hold Time after Write Strobe	50		ns
t_{wd}	Data Bus Hold Time after Write Strobe	100		ns
t_{ww}	Write Strobe Width (50% Point)	430		ns
t_{wss}	Write Strobe to Speech Output Delay		410	μs

Note: Rise and fall times (10% to 90%) of MICROBUSTM signals should be 50 ns maximum.

Timing Waveforms (Required in external processor applications)

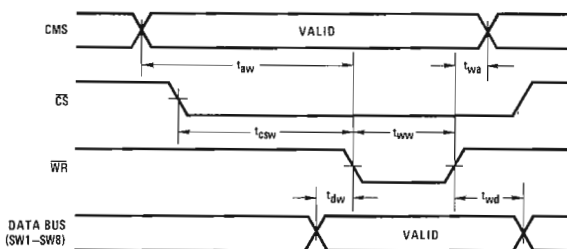
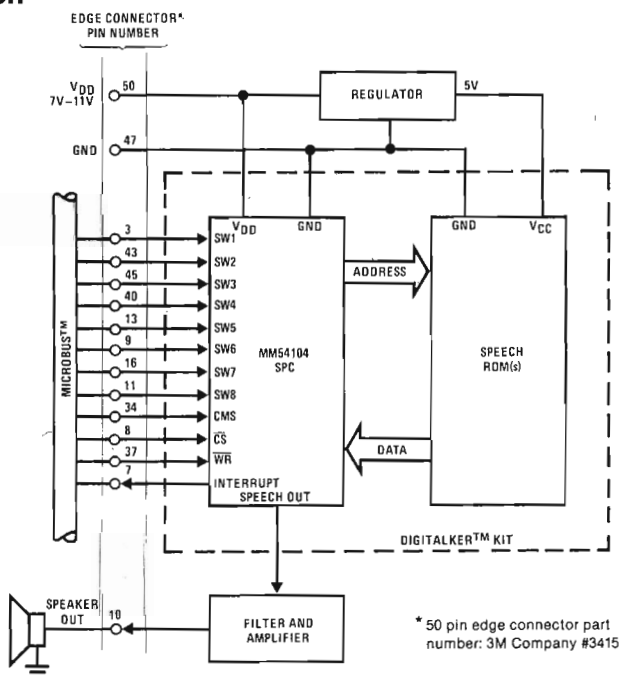


FIGURE 2. Command Sequence

Typical Application



* 50 pin edge connector part number: 3M Company #3415

Note: COP402 must be removed from DT1000 in this configuration.

FIGURE 3. DIGITALKERTTM Connections to External MICROBUSTM Processor



National Semiconductor

DIGITALKER™ Speech Synthesis

DT1050/DT1053 DIGITALKER™ Standard Vocabulary Kit

General Description

The DIGITALKER™ is a speech synthesis system consisting of several N-channel MOS integrated circuits. It contains a speech processor chip (SPC) and speech ROM and when used with external filter, amplifier, and speaker, produces a system which generates high quality speech including the natural inflection and emphasis of the original speech. Male, female, and children's voices can be synthesized.

The SPC communicates with the speech ROM, which contains the compressed speech data as well as the frequency and amplitude data required for speech output. Up to 128k bits of speech data can be directly accessed.

With the addition of an external resistor, on-chip debounce is provided for use with a switch interface.

An interrupt is generated at the end of each speech sequence so that several sequences or words can be cascaded to form different speech expressions.

The DT1050 is a standard DIGITALKER kit encoded with 137 separate and useful words, 2 tones, and 5 different silence durations. (See the Master Word List Table I). The words and tones have been assigned discrete addresses, making it possible to output single words or words concatenated into phrases or even sentences.

The "voice" output of the DT1050 is a highly intelligible male voice. The vocabulary is chosen so that it is applicable to many products and markets.

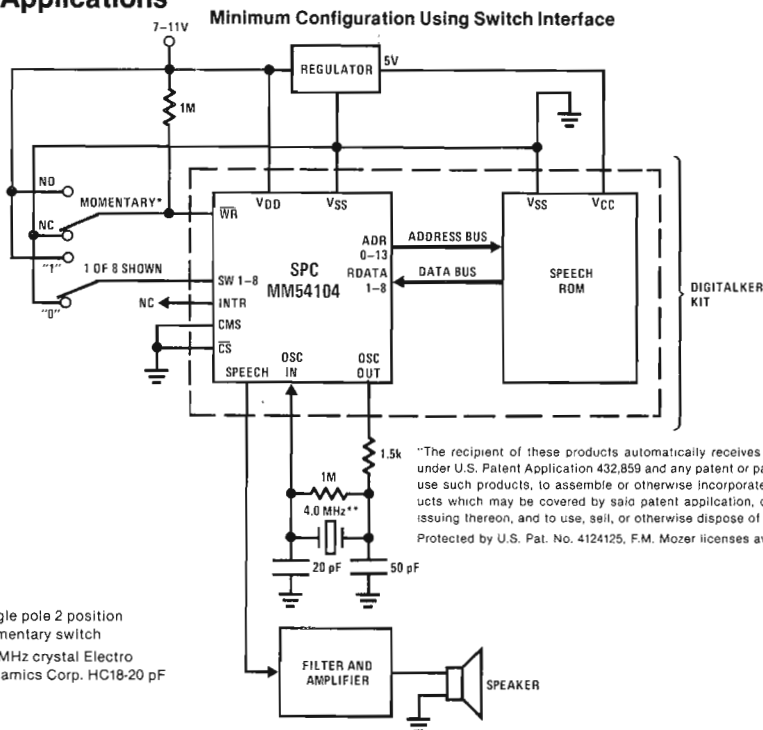
Features

- COPS™ and MICROBUS™ compatible
- Designed to be easily interfaced to other popular microprocessors
- 144 addressable expressions, including numbers
- Natural inflection and emphasis of original speech
- Addresses 128k of ROM directly
- TTL compatible
- On-chip switch debounce for interfacing to manual switches independent of a microprocessor
- Interrupt capability for cascading words or phrases
- Crystal controlled or externally driven oscillator
- Available in complete kit (DT1050) or speech ROM only (DT1053)

Applications

- Telecommunications
- Appliance
- Automotive
- Teaching aids
- Consumer products
- Clocks
- Language translation
- Annunciators

Typical Applications



DIGITALKER™, MICROBUS™ and COPS™ are trademarks of National Semiconductor Corp.

Absolute Maximum Ratings*

Storage Temperature Range	- 65°C to + 150°C	Voltage at Any Pin	12V
Operating Temperature Range	0°C to 70°C	Operating Voltage Range, $V_{DD}-V_{SS}$	7V to 11V
$V_{DD}-V_{SS}$	12V	Lead Temperature (Soldering, 10 seconds)	300°C

DC Electrical Characteristics* $T_A = 0^\circ\text{C}$ to 70°C , $V_{DD} = 7\text{V}-11\text{V}$, $V_{SS} = 0\text{V}$, unless otherwise specified.

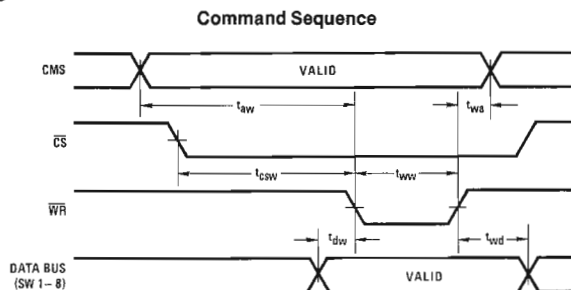
Symbol	Parameter	Conditions	Min	Typ	Max	Units
V_{IL}	Input Low Voltage		- 0.3		0.8	V
V_{IH}	Input High Voltage		2.0		V_{DD}	V
V_{OL}	Output Low Voltage	$I_{OL} = 1.6\text{ mA}$			0.4	V
V_{OH}	Output High Voltage	$I_{OH} = -100\text{ }\mu\text{A}$	2.4		5.0	V
V_{ILX}	Clock Input Low Voltage		- 0.3		1.2	V
V_{IHX}	Clock Input High Voltage		5.5		V_{DD}	V
I_{DD}	Power Supply Current				45	mA
I_{IL}	Input Leakage				± 10	μA
I_{ILX}	Clock Input Leakage				± 10	μA
V_S	Silence Voltage			0.45 V_{DD}		V
V_{OUT}	Peak to Peak Speech Output	$V_{DD} = 11\text{V}$		2.0		V
R_{EXT}	External Load on Speech Output	R_{EXT} Connected Between Speech Output and V_{SS}	50			k Ω

AC Electrical Characteristics* $T_A = 0^\circ\text{C}$ to 70°C , $V_{DD} = 7\text{V}-11\text{V}$, $V_{SS} = 0\text{V}$, unless otherwise specified.

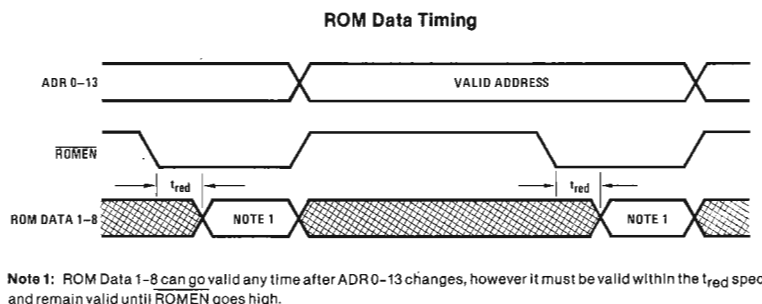
Symbol	Parameter	Min	Max	Units
t_{aw}	CMS Valid to Write Strobe	350		ns
t_{csw}	Chip Select ON to Write Strobe	310		ns
t_{dw}	Data Bus Valid to Write Strobe	50		ns
t_{wa}	CMS Hold Time after Write Strobe	50		ns
t_{wd}	Data Bus Hold Time after Write Strobe	100		ns
t_{ww}	Write Strobe Width (50% Point)	430		ns
t_{red}	ROMEN ON to Valid ROM Data		2	μs
t_{wss}	Write Strobe to Speech Output Delay		410	μs
f_t	External Clock Frequency	3.92	4.08	MHz

Note: Rise and fall times (10% to 90%) of MICROBUS signals should be 50 ns maximum.

*SPC characteristics only. ROM characteristics covered by separate data sheet for MM52164.

Timing Waveforms

Timing Waveforms (Continued)



Functional Description

The following describes the function of all SPC input and output pins.

Note: In the following descriptions, a low represents a logic 0 (0.4V nominal), and a high represents a logic 1 (2.4V nominal).

INPUT SIGNALS

Chip Select ($\overline{CS}$): The SPC is selected when $\overline{CS}$ is low. It is only necessary to have $\overline{CS}$ low during a command to the SPC. It is not necessary to hold $\overline{CS}$ low for the duration of the speech data.

Data Bus (SW 1-8): This is an 8-bit parallel data bus which contains the starting address of the speech data.

Data bus inputs SW 1-SW 8 accept an 8-bit binary address which is the address of the word which is to be "spoken" from the DIGITALKER output. See the Master Word List (Table I) for the complete listing of words and their respective addresses. If the entire word list is not used, unused inputs must be connected to V_{SS} .

Command Select (CMS): This line specifies the two commands to the SPC.

CMS	Function
0	Reset interrupt and start speech sequence
1	Reset interrupt only

Write Strobe ($\overline{WR}$): This line latches the starting address (SW 1-SW 8) into a register. On the rising edge of the $\overline{WR}$, the SPC starts execution of the command specified by CMS. The command sequence is shown in the timing waveform section. If a command to start a new speech sequence is issued during a speech sequence, the new speech sequence will be started immediately. When connecting $\overline{WR}$ to a switch, it must be a single pole 2 position switch as shown on page 1.

ROM Data (RDATA 1-8): This is an 8-bit parallel data bus which contains the speech data from the speech ROM.

OUTPUT SIGNALS

Interrupt (INTR): This signal goes high at the completion of any speech sequence. It is reset by the next valid command. It is also reset at power up.

ROM Address (ADR 0-ADR 13): This is a 14-bit parallel bus that supplies the address of the speech data to the speech ROM.

ROM Enable ($\overline{ROMEN}$): For low power applications, this line can be used to drive a transistor that switches the supply for static speech ROMs. See ROM data timing.

Speech Output (Speech Out): This is the analog output that represents the speech data. See frequency response section.

INPUT/OUTPUT SIGNALS

Clock Input/Output (OSC IN, OSC OUT): These two pins connect the main timing reference (crystal) to the SPC.

PHRASE QUALITY

In normal human speech, the brain puts durations of silence between the words to make the sentence flow smoothly. Since several durations of silence are provided in the Master Word List, the actual quality of any phrase can be significantly improved by adding durations of silence (also assigned addresses) between the words. As one thinks about how the phrase is actually spoken, one might assume the approximate duration of silence between each word, and insert the closest duration of silence from the word list. A hint in this area would be that for words beginning with the letters, K, T, P, B, D, and G insert 80 milliseconds silence prior to the words, and for words ending in the same letters as above, 40 milliseconds silence following the word is recommended.

Functional Description (Continued)

TABLE I. DT1050 MASTER WORD LIST*

Word	8-Bit Binary Address			8-Bit Binary Address			8-Bit Binary Address	
	SW 8	SW 1		SW 8	SW 1		SW 8	SW 1
THIS IS DIGITALKER	00000000		Q	00110000		IS	01100000	
ONE	00000001		R	00110001		IT	01100001	
TWO	00000010		S	00110010		KILO	01100010	
THREE	00000011		T	00110011		LEFT	01100011	
FOUR	00000100		U	00110100		LESS	01100100	
FIVE	00000101		V	00110101		LESSER	01100101	
SIX	00000110		W	00110110		LIMIT	01100110	
SEVEN	00000111		X	00110111		LOW	01100111	
EIGHT	00001000		Y	00111000		LOWER	01101000	
NINE	00001001		Z	00111001		MARK	01101001	
TEN	00001010		AGAIN	00111010		METER	01101010	
ELEVEN	00001011		AMPERE	00111011		MILE	01101011	
TWELVE	00001100		AND	00111100		MILLI	01101100	
THIRTEEN	00001101		AT	00111101		MINUS	01101101	
FOURTEEN	00001110		CANCEL	00111110		MINUTE	01101110	
FIFTEEN	00001111		CASE	00111111		NEAR	01101111	
SIXTEEN	00010000		CENT	01000000		NUMBER	01110000	
SEVENTEEN	00010001		400HERTZ TONE	01000001		OF	01110001	
EIGHTEEN	00010010		80HERTZ TONE	01000010		OFF	01110010	
NINETEEN	00010011		20MS SILENCE	01000011		ON	01110011	
TWENTY	00010100		40MS SILENCE	01000100		OUT	01110100	
THIRTY	00010101		80MS SILENCE	01000101		OVER	01110101	
FORTY	00010110		160MS SILENCE	01000110		PARENTHESIS	01110110	
FIFTY	00010111		320MS SILENCE	01000111		PERCENT	01110111	
SIXTY	00011000		CENTI	01001000		PLEASE	01111000	
SEVENTY	00011001		CHECK	01001001		PLUS	01111001	
EIGHTY	00011010		COMMA	01001010		POINT	01111010	
NINETY	00011011		CONTROL	01001011		POUND	01111011	
HUNDRED	00011100		DANGER	01001100		PULSES	01111100	
THOUSAND	00011101		DEGREE	01001101		RATE	01111101	
MILLION	00011110		DOLLAR	01001110		RE	01111110	
ZERO	00011111		DOWN	01001111		READY	01111111	
A	00100000		EQUAL	01010000		RIGHT	10000000	
B	00100001		ERROR	01010001		SS (Note 1)	10000001	
C	00100010		FEET	01010010		SECOND	10000010	
D	00100011		FLOW	01010011		SET	10000011	
E	00100100		FUEL	01010100		SPACE	10000100	
F	00100101		GALLON	01010101		SPEED	10000101	
G	00100110		GO	01010110		STAR	10000110	
H	00100111		GRAM	01010111		START	10000111	
I	00101000		GREAT	01011000		STOP	10001000	
J	00101001		GREATER	01011001		THAN	10001001	
K	00101010		HAVE	01011010		THE	10001010	
L	00101011		HIGH	01011011		TIME	10001011	
M	00101100		HIGHER	01011100		TRY	10001100	
N	00101101		HOUR	01011101		UP	10001101	
O	00101110		IN	01011110		VOLT	10001110	
P	00101111		INCHES	01011111		WEIGHT (Note 2)	10001111	

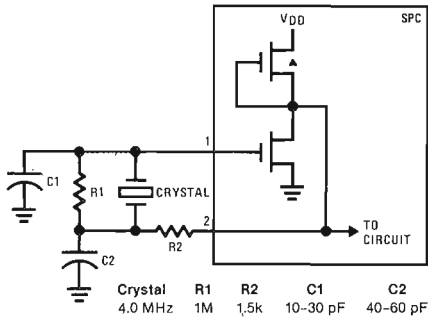
* DT1050 is a complete kit including MM54104 SPC; DT1053 is SSR1 and SSR2 speech ROMs only.

Note 1: "SS" makes any singular word plural

Note 2: Address 143 is the last legal address in this particular word list. Exceeding address 143 will produce pieces of unintelligible invalid speech data.

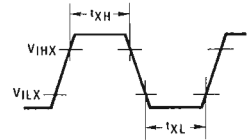
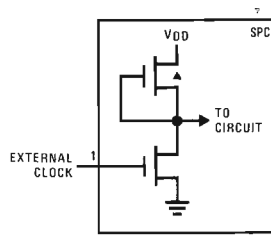
Crystal Circuit Information

Typical Crystal Oscillator Network



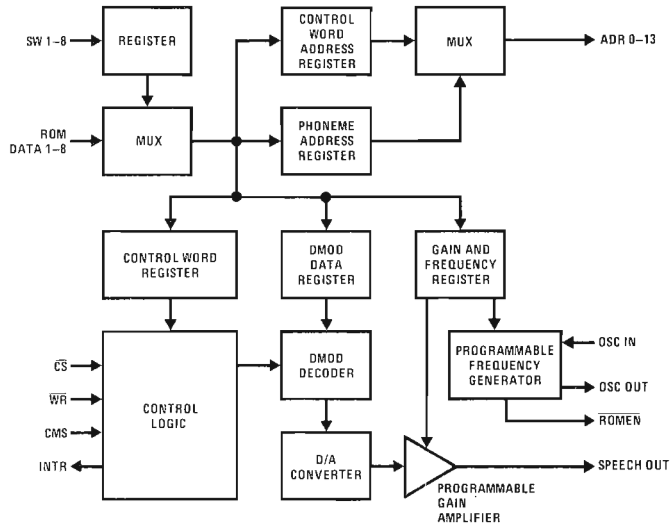
(4.0 MHz crystal manufactured by Electro Dynamics Corp. P/N HC18-20 pF)

External Clock Input (4.0 MHz)

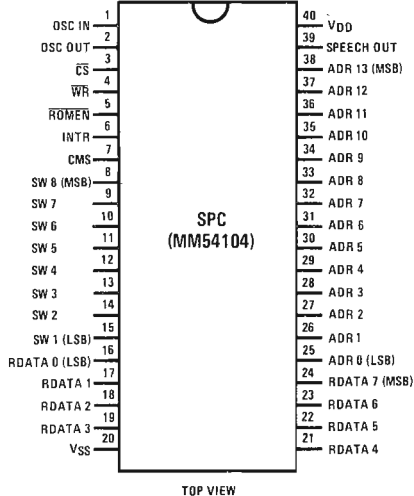


Timing	Min	Units
t_{XH}	100	ns
t_{XL}	100	ns

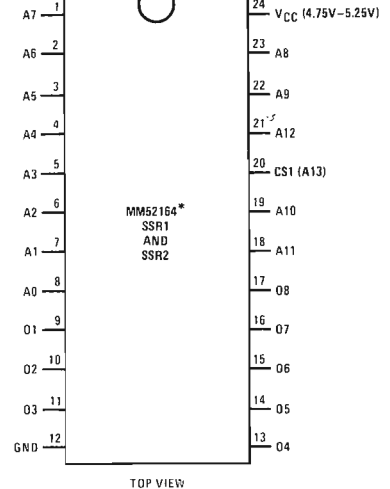
SPC Block and Connection Diagrams



Dual-In-Line Package



Dual-In-Line Package



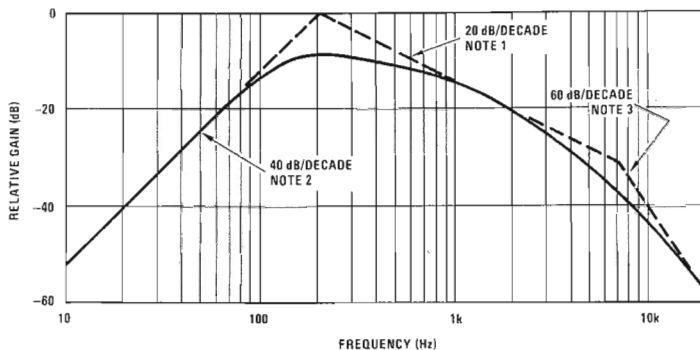
* For specific ROM device information, see MM52164 data sheet.

Scanned by the Speech Synthesizer Museum



Applications Information

Frequency Response of Combined Amplifier and Speaker



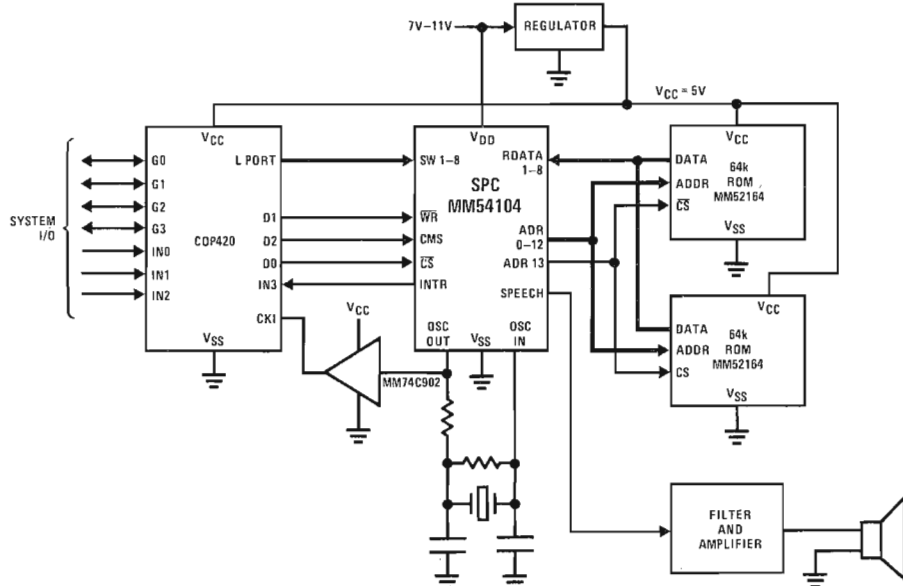
Note 1: This curve is the desired response of the entire audio system including speaker. Minimum response is a low pass filter with a cutoff frequency of 200 Hz. For an audio system with a natural cutoff frequency around 200 Hz, this filter can be eliminated. This cutoff frequency may be tuned for the particular voice being synthesized. For a low pitched male voice it may be 100 Hz, while for a high pitched female or child's voice it might be 300 Hz.

Note 2: This is optional filtering that can be eliminated by proper selection of the speaker. If this 2 pole response is electronically produced, it should be adjusted as described in Note 1.

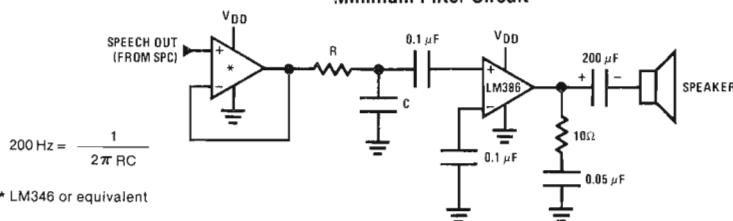
Note 3: This is optional filtering that can be eliminated for simpler systems. The acceptable range for this cutoff frequency is 6000 Hz-8000 Hz.

Typical Applications (Continued)

DIGITALKER System Using COP420 Interface

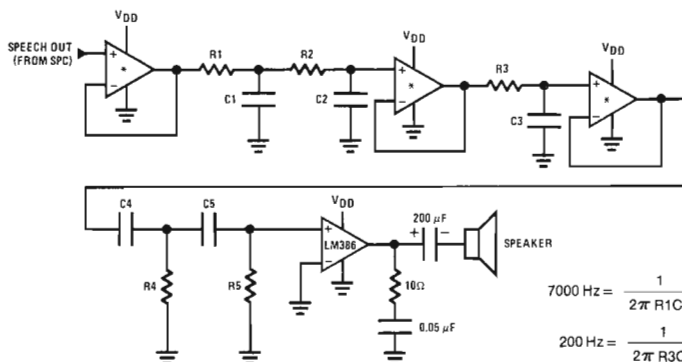


Minimum Filter Circuit



Typical Applications (Continued)

Filter Circuit to Produce Maximum Frequency Response

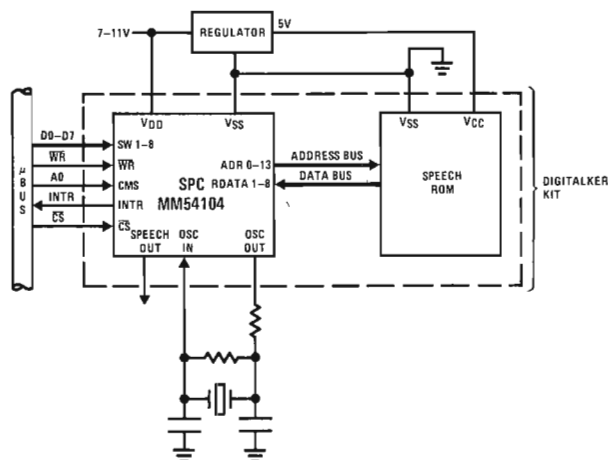


$$7000 \text{ Hz} = \frac{1}{2\pi R1C1} = \frac{1}{2\pi R2C2}$$

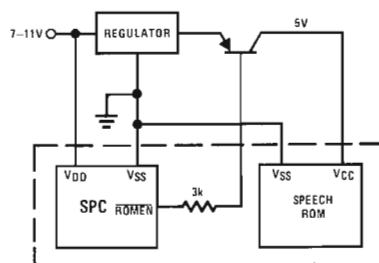
$$200 \text{ Hz} = \frac{1}{2\pi R3C3} = \frac{1}{2\pi R4C4} = \frac{1}{2\pi R5C5}$$

* LM346 or equivalent

DIGITALTALKER System Utilizing MICROBUS™ Interface



Low Power Configuration Using Static ROM




**National
Semiconductor**
DIGITALKER™ Speech Synthesis

DT1051/DT1054 DIGITALKER™ Speech Evaluation Kit

General Description

The DIGITALKER™ is a speech synthesis system consisting of several N-channel MOS integrated circuits. It contains a speech processor chip (SPC) and speech ROM and when used with external filter, amplifier, and speaker, produces a system which generates high quality speech including the natural inflection and emphasis of the original speech. Male, female, and children's voices can be synthesized.

The SPC communicates with the speech ROM, which contains the compressed speech data as well as the frequency and amplitude data required for speech output. Up to 128k bits of speech data can be directly accessed.

With the addition of an external resistor, on-chip debounce is provided for use with a switch interface.

An interrupt is generated at the end of each speech sequence so that several sequences or words can be cascaded to form different speech expressions.

The DT1051 is a standard DIGITALKER kit encoded with 18 separate and addressable phrases. The phrases lend themselves to a variety of products, but are primarily for demonstration purposes, portraying DIGITALKER's unique ability to produce high quality male, female, or child's voices.

The DT1051 demonstrates the effects of digitizing complete phrases as opposed to individual words at single

addresses. Complete phrases produce very natural and highest quality speech, much like a tape recording of a phrase or sentence.

Features

- COPST™ and MICROBUST™ compatible
- Designed to be easily interfaced to other popular microprocessors
- 18 addressable high quality phrases
- Natural inflection and emphasis of original speech
- Addresses 128k of ROM directly
- TTL compatible
- On-chip switch debounce for interfacing to manual switches independent of a microprocessor
- Interrupt capability for cascading words or phrases
- Crystal controlled or externally driven oscillator
- Available in complete kit (DT1051) or speech ROMs only (DT1054)

Applications

- Telecommunications
- Appliance
- Automotive
- Teaching aids
- Consumer products
- Clocks
- Language translation
- Annunciators

DT1051 Vocabulary*

Byte Address SW 8	SW 1	Message	Voice	Byte Address SW 8	SW 1	Message	Voice
00000000		BASSON MUSIC	N/A	00001001		CHECK OIL LEVEL	MALE
00000001		THIS IS NS DIGITALKER	FEMALE	00001010		CHECK COOLANT LEVEL	MALE
00000010		THE TIME IS 8:43 PM	FEMALE	00001011		CHECK FUEL LEVEL	MALE
00000011		EMERGENCY, CALL 911	FEMALE	00001100		DOOR OPEN	MALE
00000100		SELECT TEMPERATURE	FEMALE	00001101		DEFROST	MALE
00000101		SELECT COOKING TIME	FEMALE	00001110		GOING UP	MALE
00000110		THE NUMBER YOU REACHED HAS BEEN CHANGED, PLEASE CALL 408 737-5000	MALE	00001111		FIRST FLOOR	MALE
00000111		WARNING THE BRAKE FLUID IS LOW	MALE	00010000		PLEASE CALL YOUR OFFICE	MALE
00001000		PLEASE FASTEN YOUR SEATBELT	MALE	00010001		I'M CUTE, AREN'T I?	CHILD
						END OF VOCABULARY	

* DT1051 is a complete kit including MM54104 SPC; DT1054 is SSR3 and SSR4 speech ROMs only.

"The recipient of these products automatically receives a non-exclusive license under U.S. Patent Application 432,859 and any patent or patents issuing thereon to use such products, to assemble or otherwise incorporate them into further products which may be covered by said patent application, or any patent or patents issuing thereon, and to use, sell, or otherwise dispose of such products".

Protected by U.S. Pat. No. 4124125. F.M. Mozer licenses available.

DIGITALKER™, MICROBUST™ and COPST™ are trademarks of National Semiconductor Corp.



DT1052/DT1055 DIGITALKER™ Basic Numbers Kit

General Description

The DIGITALKER™ is a speech synthesis system consisting of several N-channel MOS integrated circuits. It contains a speech processor chip (SPC) and speech ROM and when used with external filter, amplifier, and speaker, produces a system which generates high quality speech including the natural inflection and emphasis of the original speech. Male, female, and children's voices can be synthesized.

The SPC communicates with the speech ROM, which contains the compressed speech data as well as the frequency and amplitude data required for speech output. Up to 128k bits of speech data can be directly accessed.

With the addition of an external resistor, on-chip debounce is provided for use with a switch interface.

An interrupt is generated at the end of each speech sequence so that several sequences or words can be cascaded to form different speech expressions.

The DT1052 is a useful, low cost standard DIGITALKER kit, encoded with the numbers 0 through 9, the word point, and several silence durations, located at individual addresses.

These numerical outputs can be applicable to many consumer and industrial products where numbers are frequently displayed.

Features

- COPST™ and MICROBUS™ compatible
- Designed to be easily interfaced to other popular microprocessors
- 15 addressable expressions, including silences
- Natural inflection and emphasis of original speech
- Addresses 16k of ROM directly
- TTL compatible
- On-chip switch debounce for interfacing to manual switches independent of a microprocessor
- Interrupt capability for cascading words or phrases
- Crystal controlled or externally driven oscillator
- Available in complete kit (DT1052) or speech ROM only (DT1055)*

Applications

- Telecommunications
- Appliance
- Automotive
- Counters
- Consumer products
- Instrumentation

DT1052 Vocabulary

Byte Address SW 8	SW 1	Message	Byte Address SW 8	SW 1	Message
0 0 0 0 0 0 0 0		ZERO	0 0 0 0 1 0 0 0		EIGHT
0 0 0 0 0 0 0 1		ONE	0 0 0 0 1 0 0 1		NINE
0 0 0 0 0 0 1 0		TWO	0 0 0 0 1 0 1 0		POINT
0 0 0 0 0 0 1 1		THREE	0 0 0 0 1 0 1 1		20 ms SILENCE
0 0 0 0 0 1 0 0		FOUR	0 0 0 0 1 1 0 0		40 ms SILENCE
0 0 0 0 0 1 0 1		FIVE	0 0 0 0 1 1 0 1		80 ms SILENCE
0 0 0 0 0 1 1 0		SIX	0 0 0 0 1 1 1 0		160 ms SILENCE
0 0 0 0 0 1 1 1		SEVEN	0 0 0 0 1 1 1 1		320 ms SILENCE

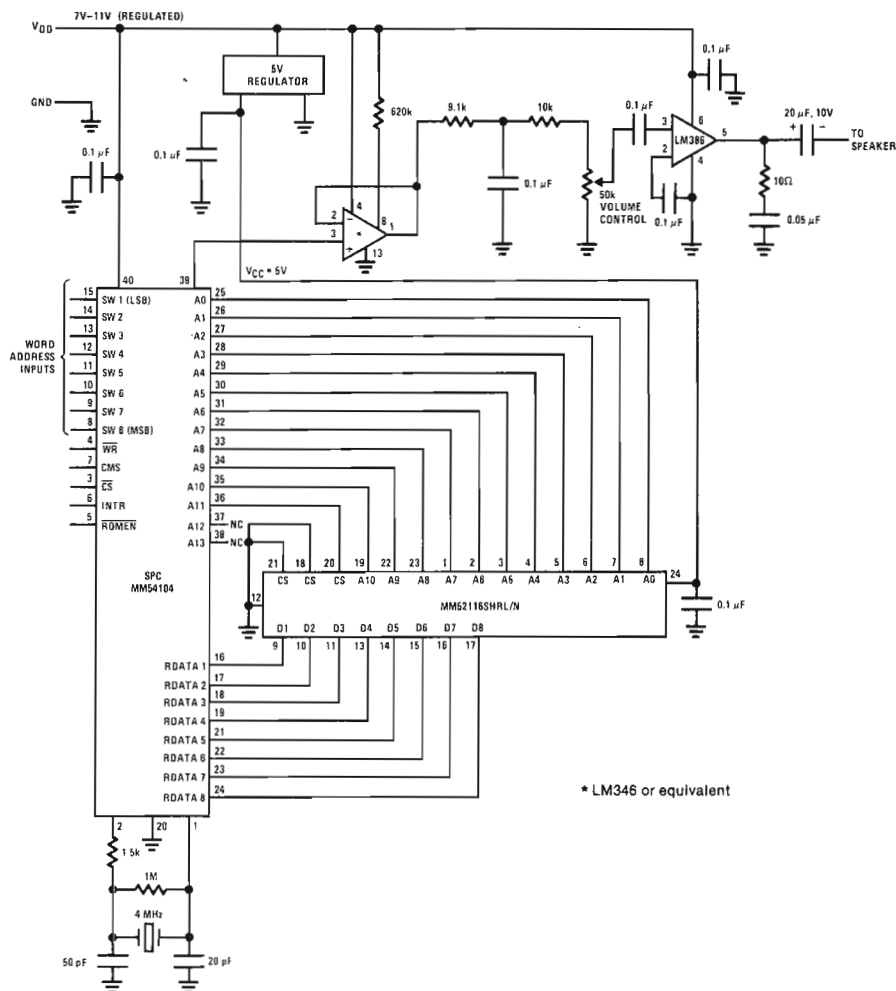
The recipient of these products automatically receives a non-exclusive license under U.S. Patent Application 432,859 and any patent or patents issuing thereon to use such products, to assemble or otherwise incorporate them into further products which may be covered by said patent application, or any patent or patents issuing thereon, and to use, sell, or otherwise dispose of such products.

Protected by U.S. Pat. No. 4124125, F.M. Mozer licenses available.

* DT1052 is a complete kit including MM54104 SPC; DT1055 is MM52116SHRL speech ROM only.

DIGITALKER™, MICROBUS™ and COPST™ are trademarks of National Semiconductor Corp.

Recommended Schematic Diagram



Refer to MM54104 and MM52116 data sheets for complete specifications on electrical and timing characteristics.



National Semiconductor

DIGITALKER™ Speech Synthesis

DT1056/DT1057 DIGITALKER™ Standard Vocabulary Kit

General Description

The DIGITALKER™ is a speech synthesis system consisting of several N-channel MOS integrated circuits. It contains a speech processor chip (SPC) and speech ROM and when used with external filter, amplifier, and speaker, produces a system which generates high quality speech including the natural inflection and emphasis of the original speech. Male, female, and children's voices can be synthesized.

The SPC communicates with the speech ROM, which contains the compressed speech data as well as the frequency and amplitude data required for speech output. Up to 128k bits of speech data can be directly accessed.

With the addition of an external resistor, on-chip debounce is provided for use with a switch interface.

An interrupt is generated at the end of each speech sequence so that several sequences or words can be cascaded to form different speech expressions.

The DT1056/DT1057 is a standard DIGITALKER kit encoded with 131 separate and useful words (see the Master Word List Table I) and when used with the DT1050 Standard Vocabulary Kit, provides a library of 274 useful words. The words have been assigned discrete addresses, making it possible to output single words or words concatenated into phrases or even sentences.

The "voice" output of the DT1056/DT1057 is a highly intelligible male voice. The vocabulary is chosen so that it is applicable to many products and markets.

Features

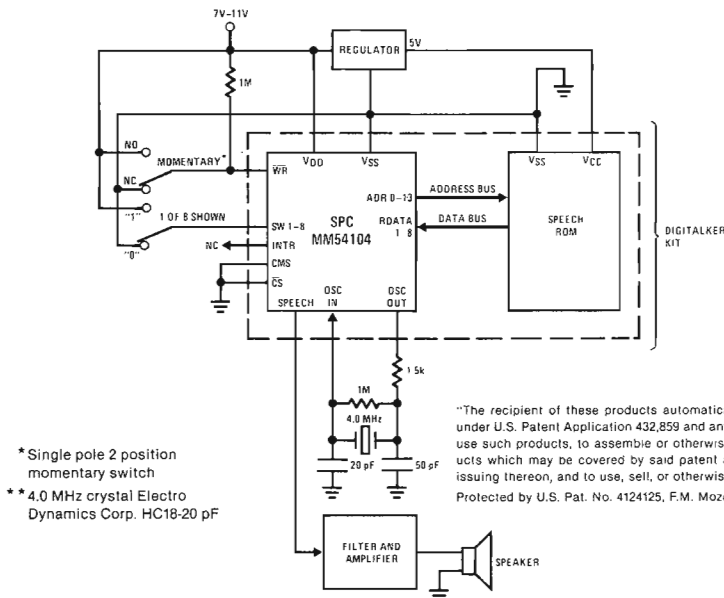
- Easily adaptable to DT1050 Standard Vocabulary Kit
- 131 useful words
- COPS™ and MICROBUS™ compatible
- Designed to be easily interfaced to other popular microprocessors
- Natural inflection and emphasis of original speech
- Addresses 128k bits of ROM directly
- TTL compatible
- On-chip switch debounce for interfacing to manual switches independent of a microprocessor
- Interrupt capability for cascading words or phrases
- Crystal controlled or externally driven oscillator
- Available in complete kit (DT1056) or speech ROMs only (DT1057)

Applications

- Telecommunications
- Appliance
- Automotive
- Teaching aids
- Consumer products
- Clocks
- Language translation
- Annunciators

Typical Applications

Minimum Configuration Using Switch Interface



"The recipient of these products automatically receives a non-exclusive license under U.S. Patent Application 432,859 and any patent or patents issuing thereon to use such products, to assemble or otherwise incorporate them into further products which may be covered by said patent application, or any patent or patents issuing thereon, and to use, sell, or otherwise dispose of such products."

Protected by U.S. Pat. No. 4124125, F.M. Mozer licenses available.

Absolute Maximum Ratings*

Storage Temperature Range	- 65°C to + 150°C
Operating Temperature Range	0°C to 70°C
V _{DD} -V _{SS}	12V

Voltage at Any Pin	12V
Operating Voltage Range, V _{DD} -V _{SS}	7V to 11V
Lead Temperature (Soldering, 10 seconds)	300°C

DC Electrical Characteristics* T_A = 0°C to 70°C, V_{DD} = 7V-11V, V_{SS} = 0V, unless otherwise specified.

Symbol	Parameter	Conditions	Min	Typ	Max	Units
V _{IL}	Input Low Voltage		- 0.3		0.8	V
V _{IH}	Input High Voltage		2.0		V _{DD}	V
V _{OL}	Output Low Voltage	I _{OL} = 1.6 mA			0.4	V
V _{OH}	Output High Voltage	I _{OH} = - 100 μ A	2.4		5.0	V
V _{ILX}	Clock Input Low Voltage		- 0.3		1.2	V
V _{IHX}	Clock Input High Voltage		5.5		V _{DD}	V
I _{DD}	Power Supply Current				45	mA
I _{IL}	Input Leakage				$\pm$ 10	μ A
I _{ILX}	Clock Input Leakage				$\pm$ 10	μ A
V _S	Silence Voltage			0.45 V _{DD}		V
V _{OUT}	Peak to Peak Speech Output	V _{DD} = 11V		2.0		V
R _{EXT}	External Load on Speech Output	R _{EXT} Connected Between Speech Output and V _{SS}	50			k Ω

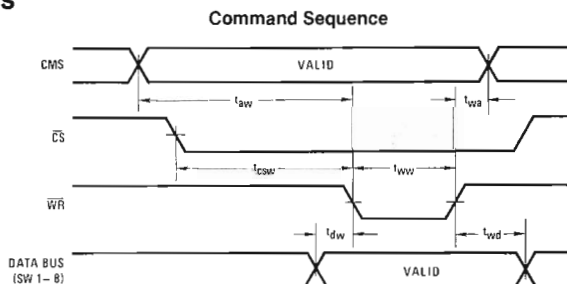
AC Electrical Characteristics* T_A = 0°C to 70°C, V_{DD} = 7V-11V, V_{SS} = 0V, unless otherwise specified.

Symbol	Parameter	Min	Max	Units
t _{aw}	CMS Valid to Write Strobe	350		ns
t _{csw}	Chip Select ON to Write Strobe	310		ns
t _{dw}	Data Bus Valid to Write Strobe	50		ns
t _{wa}	CMS Hold Time after Write Strobe	50		ns
t _{wd}	Data Bus Hold Time after Write Strobe	100		ns
t _{ww}	Write Strobe Width (50% Point)	430		ns
t _{red}	ROMEN ON to Valid ROM Data		2	μ s
t _{wss}	Write Strobe to Speech Output Delay		410	μ s
f _t	External Clock Frequency	3.92	4.08	MHz

Note: Rise and fall times (10% to 90%) of MICROBUS signals should be 50 ns maximum.

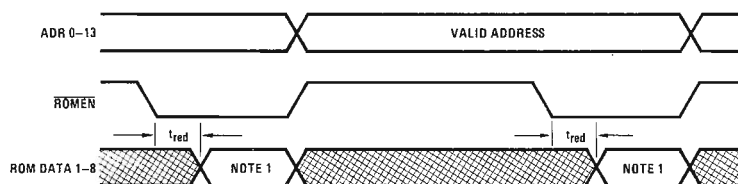
*SPC characteristics only. ROM characteristics covered by separate data sheet for MM52164.

Timing Waveforms



Timing Waveforms.(Continued)

ROM Data Timing



Note 1: ROM Data 1-8 can go valid any time after ADR 0-13 changes, however it must be valid within the t_{red} spec and remain valid until $\overline{ROMEN}$ goes high.

Functional Description

The following describes the function of all SPC input and output pins.

Note: In the following descriptions, a low represents a logic 0 (0.4V nominal), and a high represents a logic 1 (2.4V nominal).

INPUT SIGNALS

Chip Select ($\overline{CS}$): The SPC is selected when $\overline{CS}$ is low. It is only necessary to have $\overline{CS}$ low during a command to the SPC. It is not necessary to hold $\overline{CS}$ low for the duration of the speech data.

Data Bus (SW 1-8): This is an 8-bit parallel data bus which contains the starting address of the speech data.

Data bus inputs SW 1-SW 8 accept an 8-bit binary address which is the address of the word which is to be "spoken" from the DIGITALKER output. See the Master Word List (Table I) for the complete listing of words and their respective addresses. If the entire word list is not used, unused inputs must be connected to V_{SS} .

Command Select (CMS): This line specifies the two commands to the SPC.

CMS	Function
0	Reset interrupt and start speech sequence
1	Reset interrupt only

Write Strobe ($\overline{WR}$): This line latches the starting address (SW 1-SW 8) into a register. On the rising edge of the $\overline{WR}$, the SPC starts execution of the command specified by CMS. The command sequence is shown in the timing waveform section. If a command to start a new speech sequence is issued during a speech sequence, the new speech sequence will be started immediately. When connecting $\overline{WR}$ to a switch, it must be a single pole 2 position switch as shown on page 1.

ROM Data (RDATA 1-8): This is an 8-bit parallel data bus which contains the speech data from the speech ROM.

OUTPUT SIGNALS

Interrupt (INTR): This signal goes high at the completion of any speech sequence. It is reset by the next valid command. It is also reset at power up.

ROM Address (ADR 0-ADR 13): This is a 14-bit parallel bus that supplies the address of the speech data to the speech ROM.

ROM Enable ($\overline{ROMEN}$): For low power applications, this line can be used to drive a transistor that switches the supply for static speech ROMs. See ROM Data Timing.

Speech Output (Speech Out): This is the analog output that represents the speech data. See frequency response section.

INPUT/OUTPUT SIGNALS

Clock Input/Output (OSC IN, OSC OUT): These two pins connect the main timing reference (crystal) to the SPC.

PHRASE QUALITY

In normal human speech, the brain puts durations of silence between the words to make the sentence flow smoothly. Since several durations of silence are provided in the Master Word List, the actual quality of any phrase can be significantly improved by adding durations of silence (also assigned addresses) between the words. As one thinks about how the phrase is actually spoken, one might assume the approximate duration of silence between each word, and insert the closest duration of silence from the word list. A hint in this area would be that for words beginning with the letters, K, T, P, B, D, and G insert 80 milliseconds silence prior to the words, and for words ending in the same letters as above, 40 milliseconds silence following the word is recommended.

Functional Description (Continued)

TABLE I. DT1056/DT1057* MASTER WORD LIST

Word	8-Bit Binary Address		Word	8-Bit Binary Address		Word	8-Bit Binary Address	
	SW8	SW1		SW8	SW1		SW8	SW1
ABORT	00000000		FARAD	00101100		PER	01011000	
ADD	00000001		FAST	00101101		PICO	01011001	
ADJUST	00000010		FASTER	00101110		PLACE	01011010	
ALARM	00000011		FIFTH	00101111		PRESS	01011011	
ALERT	00000100		FIRE	00110000		PRESSURE	01011100	
ALL	00000101		FIRST	00110001		QUARTER	01011101	
ASK	00000110		FLOOR	00110010		RANGE	01011110	
ASSISTANCE	00000111		FORWARD	00110011		REACH	01011111	
ATTENTION	00001000		FROM	00110100		RECEIVE	01100000	
BRAKE	00001001		GAS	00110101		RECORD	01100001	
BUTTON	00001010		GET	00110110		REPLACE	01100010	
BUY	00001011		GOING	00110111		REVERSE	01100011	
CALL	00001100		HALF	00111000		ROOM	01100100	
CAUTION	00001101		HELLO	00111001		SAFE	01100101	
CHANGE	00001110		HELP	00111010		SECURE	01100110	
CIRCUIT	00001111		HERTZ	00111011		SELECT	01100111	
CLEAR	00010000		HOLD	00111100		SEND	01101000	
CLOSE	00010001		INCORRECT	00111101		SERVICE	01101001	
COMPLETE	00010010		INCREASE	00111110		SIDE	01101010	
CONNECT	00010011		INTRUDER	00111111		SLOW	01101011	
CONTINUE	00010100		JUST	01000000		SLOWER	01101100	
COPY	00010101		KEY	01000001		SMOKE	01101101	
CORRECT	00010110		LEVEL	01000010		SOUTH	01101110	
DATE	00010111		LOAD	01000011		STATION	01101111	
DAY	00011000		LOCK	01000100		SWITCH	01110000	
DECREASE	00011001		MEG	01000101		SYSTEM	01110001	
DEPOSIT	00011010		MEGA	01000110		TEST	01110010	
DIAL	00011011		MICRO	01000111		TH (NOTE 2)	01110011	
DIVIDE	00011100		MORE	01001000		THANK	01110100	
DOOR	00011101		MOVE	01001001		THIRD	01110101	
EAST	00011110		NANO	01001010		THIS	01110110	
ED (NOTE 1)	00011111		NEED	01001011		TOTAL	01110111	
ED (NOTE 1)	00100000		NEXT	01001100		TURN	01111000	
ED (NOTE 1)	00100001		NO	01001101		USE	01111001	
ED (NOTE 1)	00100010		NORMAL	01001110		UTH (NOTE 3)	01111010	
EMERGENCY	00100011		NORTH	01001111		WAITING	01111011	
END	00100100		NOT	01010000		WARNING	01111100	
ENTER	00100101		NOTICE	01010001		WATER	01111101	
ENTRY	00100110		OHMS	01010010		WEST	01111110	
ER	00100111		ONWARD	01010011		SWITCH	01111111	
EVACUATE	00101000		OPEN	01010100		WINDOW	10000000	
EXIT	00101001		OPERATOR	01010101		YES	10000001	
FAIL	00101010		OR	01010110		ZONE	10000010	
FAILURE	00101011		PASS	01010111				

* DT1056 is a complete kit including MM54104 SPC; DT1057 is SSR5 and SSR6 speech ROMs only.

Note 1: "ED" is a suffix that can be used to make any present tense word become a past tense word. The way we say "ED," however, does vary from one word to the next. For that reason, we have offered 4 different "ED" sounds. It is suggested that each "ED" be tested with the desired word for best quality results. Address 31 "ED" or 32 "ED" should be used with words ending in "T" or "D," such as exit or load. Address 34 "ED" should be used with words ending with soft sounds such as ask. Address 33 "ED" should be used with all other words.

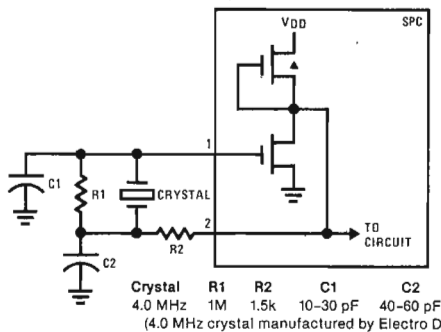
Note 2: "TH" is a suffix that can be added to words like six, seven, eight to form adjective words like sixth, seventh, eighth.

Note 3: "UTH" is a suffix that can be added to words like twenty, thirty, forty to form adjective words like thirtieth, fortieth, etc.

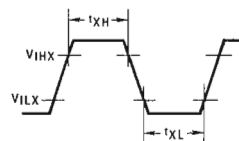
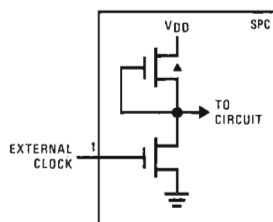
Note 4: Address 130 is the last legal address in this particular word list. Exceeding address 130 will produce pieces of unintelligible invalid speech data.

Crystal Circuit Information

Typical Crystal Oscillator Network

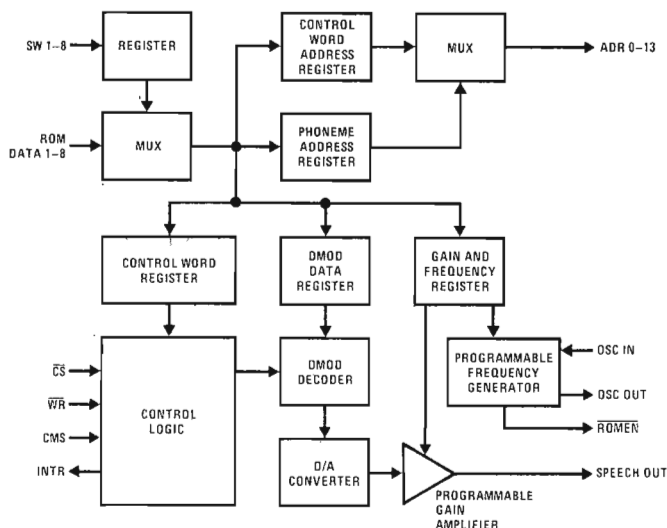


External Clock Input (4.0 MHz)

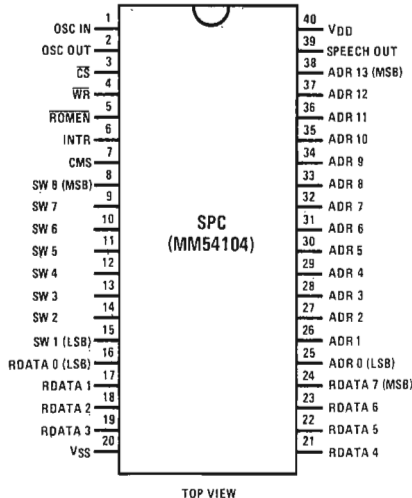


Timing	Min	Units
t_{XH}	100	ns
t_{XL}	100	ns

SPC Block and Connection Diagrams

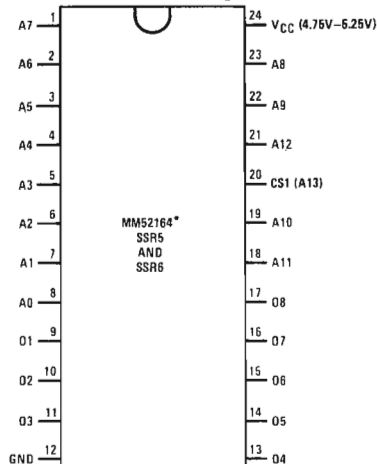


Dual-In-Line Package



TOP VIEW

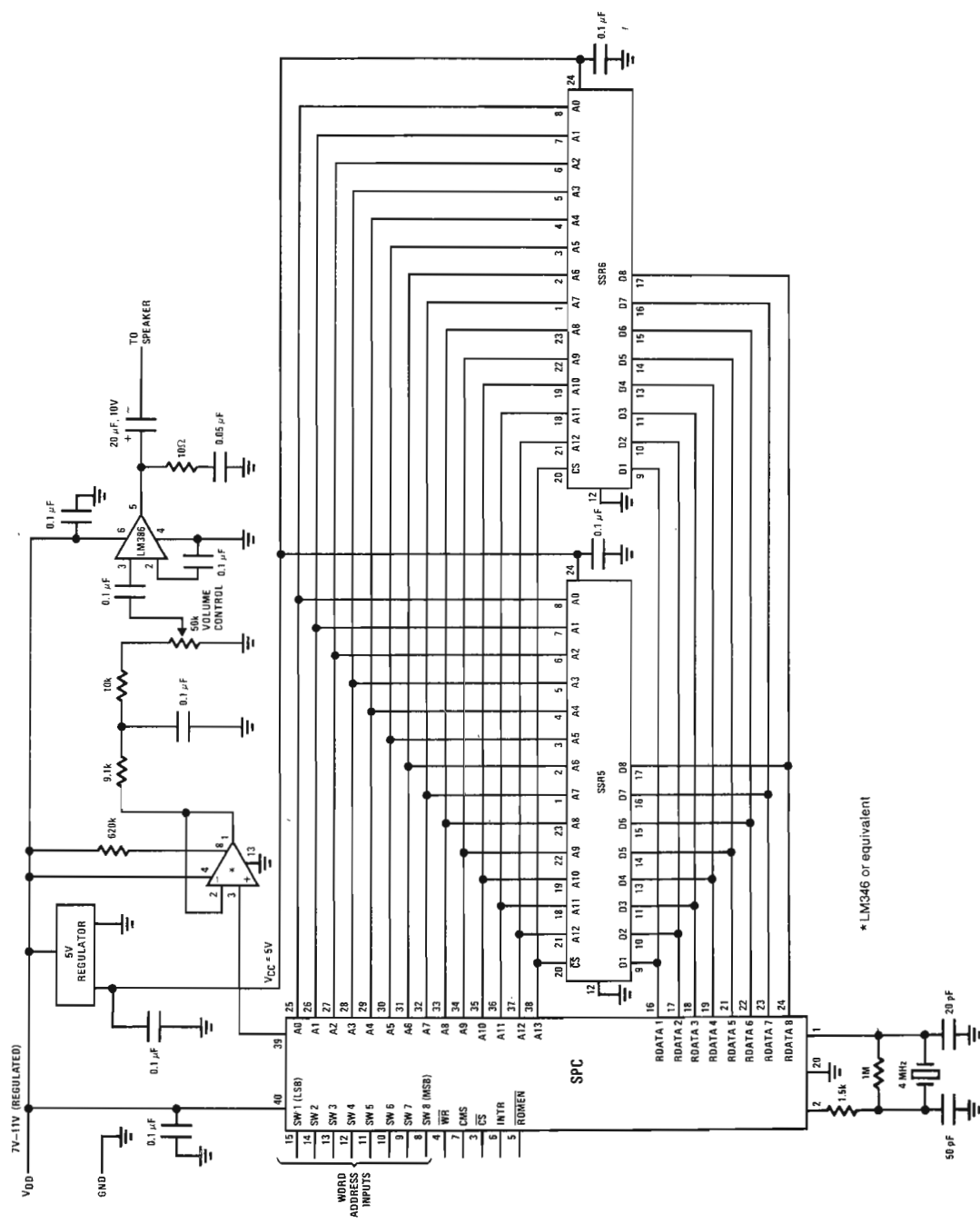
Dual-In-Line Package



TOP VIEW

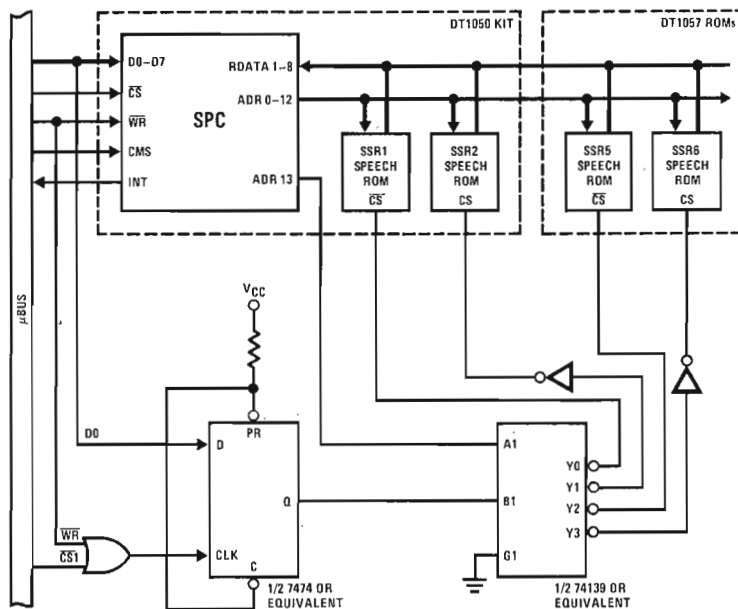
* For specific ROM device information, see MM52164 data sheet.

Recommended Schematic Diagram

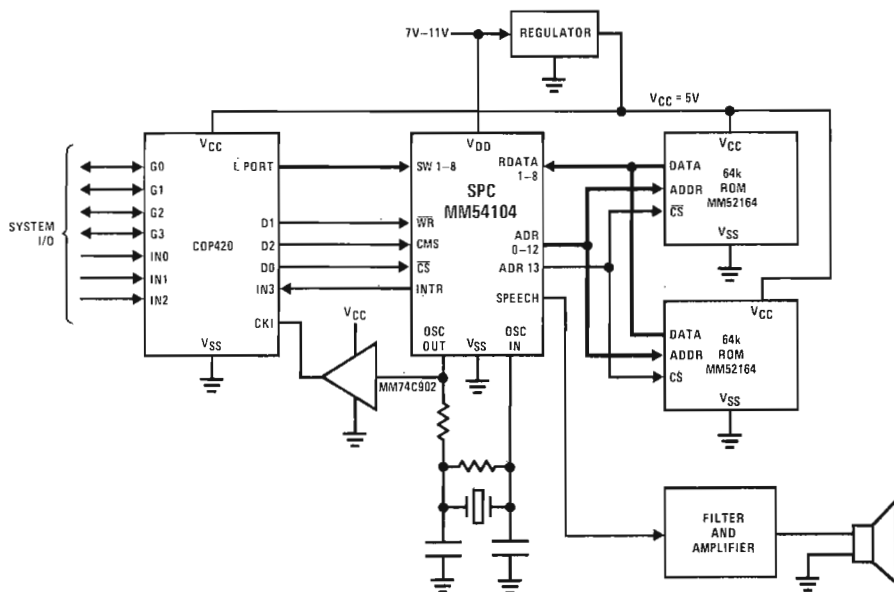


Typical Applications (Continued)

Integration of DT1057 ROMs and DT1050 Kit

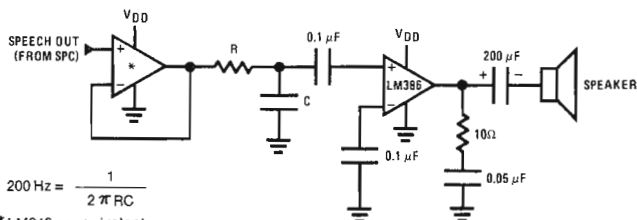


DIGITALKER System Using COP420 Interface



Typical Applications (Continued)

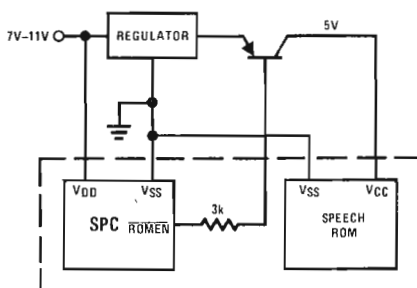
Minimum Filter Circuit



* LM346 or equivalent

See MM54104 data sheet for additional filter information.

Low Power Configuration Using Static ROM





DIGITALKER™ Speech Synthesis

MM54104 DIGITALKER™ Speech Synthesis System

General Description

The DIGITALKER is a speech synthesis system consisting of multiple N-channel MOS integrated circuits. It contains an MM54104 speech processor chip (SPC) and speech ROM and when used with external filter, amplifier, and speaker, produces a system which generates high quality speech including the natural inflection and emphasis of the original speech. Male, female, and children's voices can be synthesized.

The SPC communicates with the speech ROM, which contains the compressed speech data as well as the frequency and amplitude data required for speech output. Up to 128k bits of speech data can be directly accessed. This can be expanded with minimal external logic.

With the addition of an external resistor, on-chip debounce is provided for use with a switch interface.

An interrupt is generated at the end of each speech sequence so that several sequences or words can be cascaded to form different speech expressions.

Encoding (digitizing) of custom word or phrase lists must be done by National Semiconductor. Customers submit to the factory high quality recorded magnetic reel to reel tapes containing the words or phrases to be encoded. National Semiconductor will sell kits consisting of the SPC and ROM(s) containing the digitized word or phrases.

Features

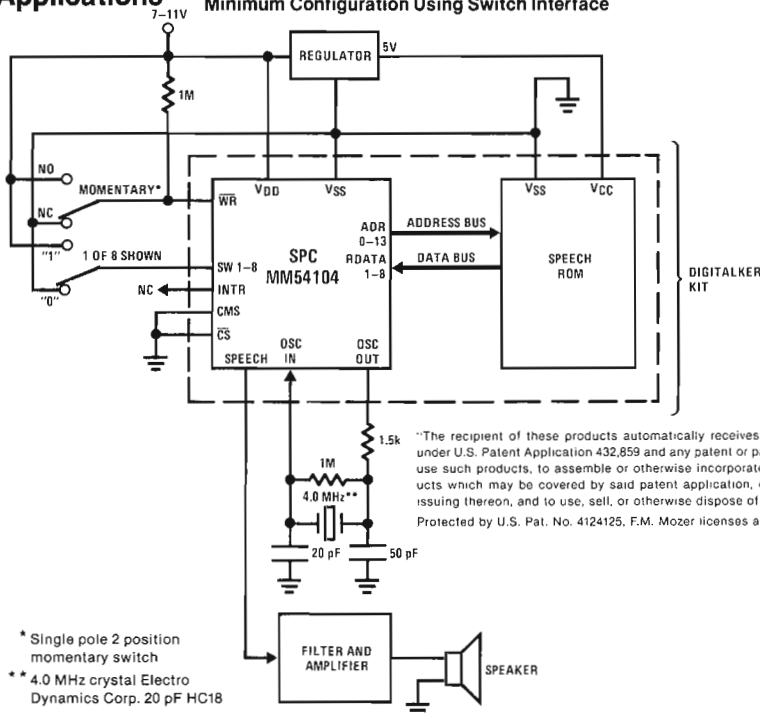
- Designed to be easily interfaced to most popular microprocessors
- 256 possible addressable expressions
- Male, female, and children's voices
- Any language
- Natural inflection and emphasis of original speech
- Addresses 128k of ROM directly
- TTL compatible
- MICROBUS™ and COPS™ compatible
- On-chip switch debounce for interfacing to manual switches independent of a microprocessor
- Easily expandable to greater than 128k ROM
- Interrupt capability for cascading words or phrases
- Crystal controlled or externally driven oscillator
- Ability to store silence durations for timing sequences
- Standard vocabulary sets available

Applications

- | | |
|----------------------|------------------------|
| ■ Telecommunications | ■ Consumer products |
| ■ Appliance | ■ Clocks |
| ■ Automotive | ■ Language translation |
| ■ Teaching aids | ■ Annunciators |

Typical Applications

Minimum Configuration Using Switch Interface



* Single pole 2 position momentary switch
 ** 4.0 MHz crystal Electro Dynamics Corp. 20 pF HC18

"The recipient of these products automatically receives a non-exclusive license under U.S. Patent Application 432,859 and any patent or patents issuing thereon to use such products, to assemble or otherwise incorporate them into further products which may be covered by said patent application, or any patent or patents issuing thereon, and to use, sell, or otherwise dispose of such products".
 Protected by U.S. Pat. No. 4124125, F.M. Mozer licenses available.

Absolute Maximum Ratings

Storage Temperature Range	- 65°C to + 150°C	Voltage at Any Pin	12V
Operating Temperature Range	- 40°C to 85°C	Operating Voltage Range, $V_{DD}-V_{SS}$	7V to 11V
$V_{DD}-V_{SS}$	12V	Lead Temperature (Soldering, 10 seconds)	300°C

DC Electrical Characteristics $T_A = 0^\circ\text{C}$ to 70°C , $V_{DD} = 7\text{V}-11\text{V}$, $V_{SS} = 0\text{V}$, unless otherwise specified.

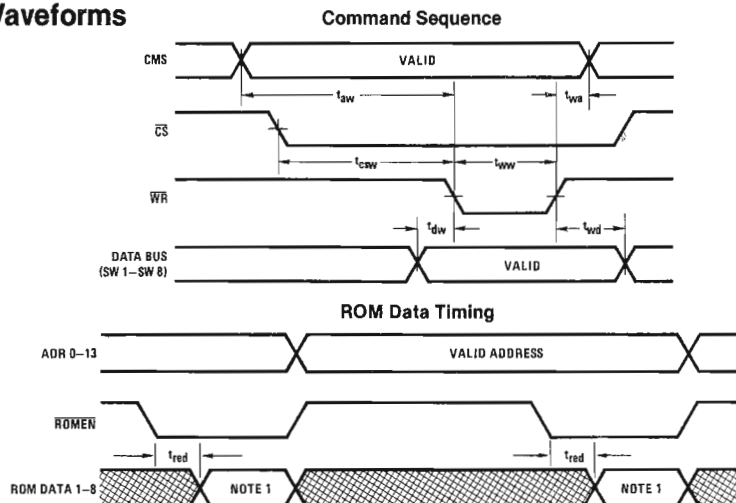
Symbol	Parameter	Conditions	Min	Typ	Max	Units
V_{IL}	Input Low Voltage	$T_A = -40^\circ\text{C}$ to 85°C	- 0.3		0.8	V
V_{IH}	Input High Voltage		- 0.3		0.6	V
V_{OL}	Output Low Voltage	$T_A = -40^\circ\text{C}$ to 85°C	2.0		V_{DD}	V
V_{OH}	Output High Voltage		2.2		V_{DD}	V
V_{ILX}	Clock Input Low Voltage	$I_{OL} = 1.6\text{ mA}$			0.4	V
V_{IHx}	Clock Input High Voltage				5.0	V
V_{OLX}	Clock Output Low Voltage	$I_{OH} = -100\text{ }\mu\text{A}$	- 0.3		1.2	V
V_{OHX}	Clock Output High Voltage		5.5		V_{DD}	V
I_{DD}	Power Supply Current	Typical Crystal Configuration and 10M Load on Pin 2			1.2	V
I_{DD}	Power Supply Current					V
I_{IL}	Input Leakage	Typical Crystal Configuration and 10M Load on Pin 2			5.5	V
I_{ILX}	Clock Input Leakage					V
V_S	Silence Voltage	Typical Crystal Configuration and 10M Load on Pin 2				V
V_{OUT}	Peak to Peak Speech Output					V
R_{EXT}	External Load on Speech Output	$V_{DD} = 11\text{V}$				k Ω
						k Ω

AC Electrical Characteristics $T_A = 0^\circ\text{C}$ to 70°C , $V_{DD} = 7\text{V}-11\text{V}$, $V_{SS} = 0\text{V}$, unless otherwise specified.

Symbol	Parameter	Min	Max	Units
t_{aw}	CMS Valid to Write Strobe	350		ns
t_{csw}	Chip Select ON to Write Strobe	310		ns
t_{dw}	Data Bus Valid to Write Strobe	50		ns
t_{wa}	CMS Hold Time after Write Strobe	50		ns
t_{wd}	Data Bus Hold Time after Write Strobe	100		ns
t_{ww}	Write Strobe Width (50% Point)	430		ns
t_{red}	ROMEN ON to Valid ROM Data		2	μs
t_{wss}	Write Strobe to Speech Output Delay		410	μs
f_t	External Clock Frequency	3.92	4.08	MHz

Note: Rise and fall times (10% to 90%) of MICROBUS signals should be 50 ns maximum.

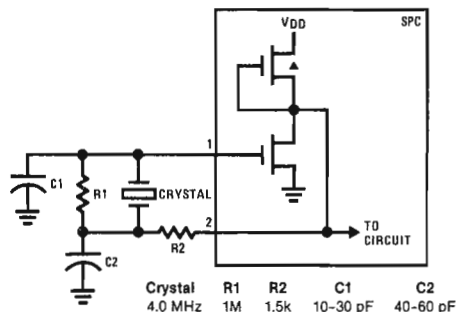
Timing Waveforms



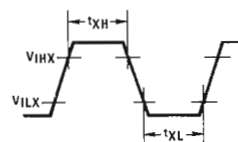
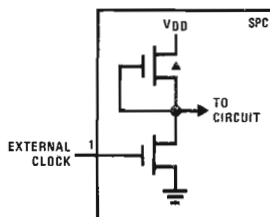
Note 1: ROM data 1-8 can go valid any time after ADR 0-13 changes, however it must be valid within the t_{red} specifications and remain valid until $\overline{ROMEN}$ goes high.

Crystal Circuit Information

Typical Crystal Oscillator Network

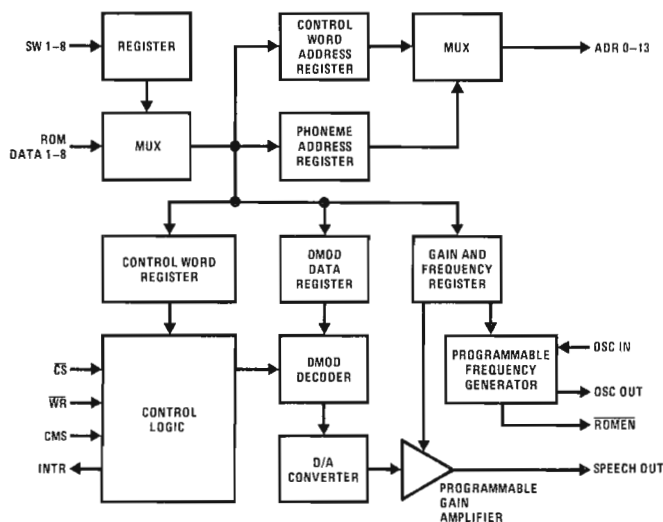


External Clock Input (4.0 MHz)

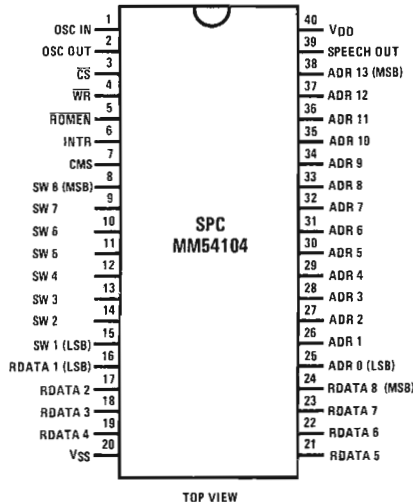


Timing	Min	Units
t_{XH}	100	ns
t_{XL}	100	ns

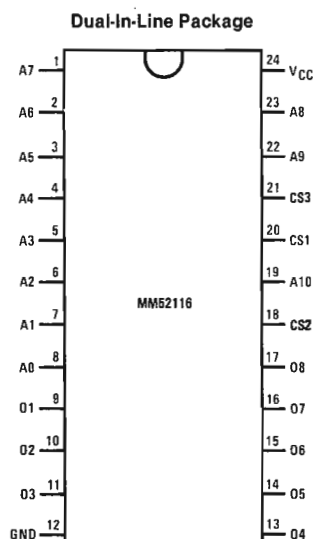
Block and Connection Diagrams



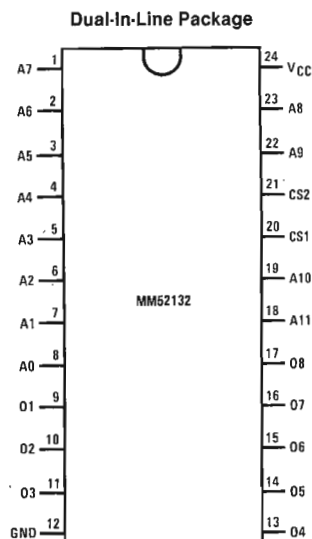
Dual-In-Line Package



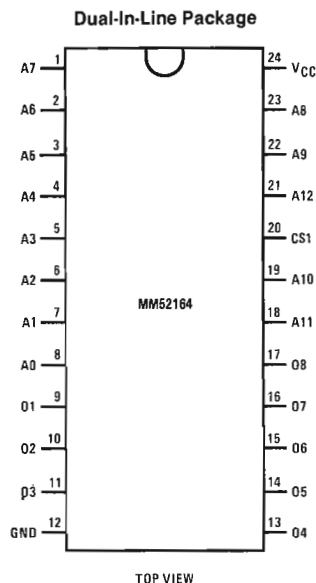
Connection Diagrams (Continued) ($V_{CC} = 4.75V-5.25V$)



16k



32k



64k

Functional Description

The following describes the function of all SPC input and output pins.

Note: In the following descriptions, a low represents a logic 0 (0.4V nominal), and a high represents a logic 1 (2.4V nominal).

INPUT SIGNALS

Chip Select ($\overline{CS}$): The SPC is selected when $\overline{CS}$ is low. It is only necessary to have $\overline{CS}$ low during a command to the SPC. It is not necessary to hold $\overline{CS}$ low for the duration of the speech data.

Data Bus (SW 1-8): This is an 8-bit parallel data bus which contains the starting address of the speech data. Unused inputs must be tied to V_{SS} .

Command Select (CMS): This line specifies the two commands to the SPC.

CMS	Function
0	Reset interrupt and start speech sequence
1	Reset interrupt only

Write Strobe ($\overline{WR}$): This line latches the starting address (SW1-SW8) into a register. On the rising edge of the $\overline{WR}$, the SPC starts execution of the command specified by CMS. The command sequence is shown in the timing waveform section. If a command to start a new speech se-

quence is issued during a speech sequence, the new speech sequence will be started immediately. When connecting $\overline{WR}$ to a switch it must be a single pole 2 position switch as shown on page 1.

ROM Data (RDATA 1-8): This is an 8-bit parallel data bus which contains the speech data from the speech ROM.

OUTPUT SIGNALS

Interrupt (INTR): This signal goes high at the completion of any speech sequence. It is reset by the next valid command. It is also reset at power up.

ROM Address (ADR 0-ADR 13): This is a 14-bit parallel bus that supplies the address of the speech data to the speech ROM.

ROM Enable ($\overline{ROMEN}$): For low power applications, this line can be used to drive a transistor that switches the supply for static speech ROMs. See ROM data timing.

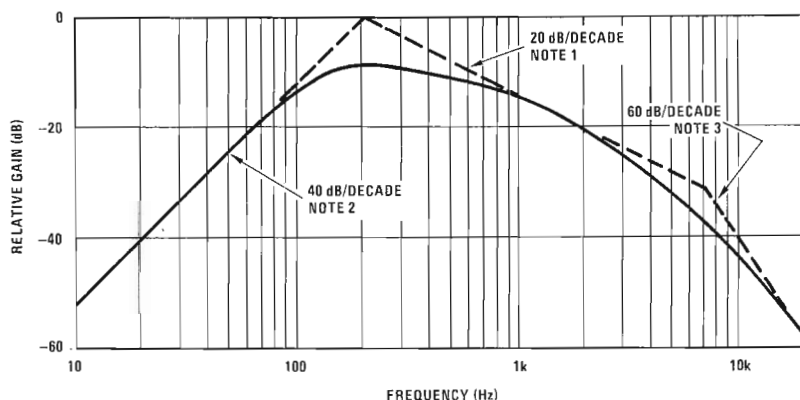
Speech Output (Speech Out): This is the analog output that represents the speech data. See frequency response section.

INPUT/OUTPUT SIGNALS

Clock Input/Output (OSC IN, OSC OUT): These two pins connect the main timing reference (crystal) to the SPC.

Applications Information

Frequency Response of Combined Amplifier and Speaker



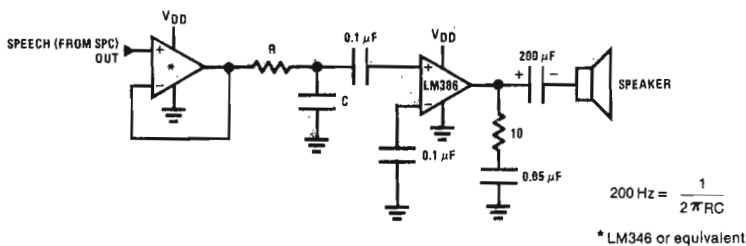
Note 1: This curve is the desired response of the entire audio system including speaker. Minimum response is a low pass filter with a cutoff frequency of 200 Hz. For an audio system with a natural cutoff frequency around 200 Hz, this filter can be eliminated. This cutoff frequency may be tuned for the particular voice being synthesized. For a low pitched male voice it may be 100 Hz, while for a high pitched female or child's voice it might be 300 Hz.

Note 2: This is optional filtering that can be eliminated by proper selection of the speaker. If this 2 pole response is electronically produced, it should be adjusted as described in Note 1.

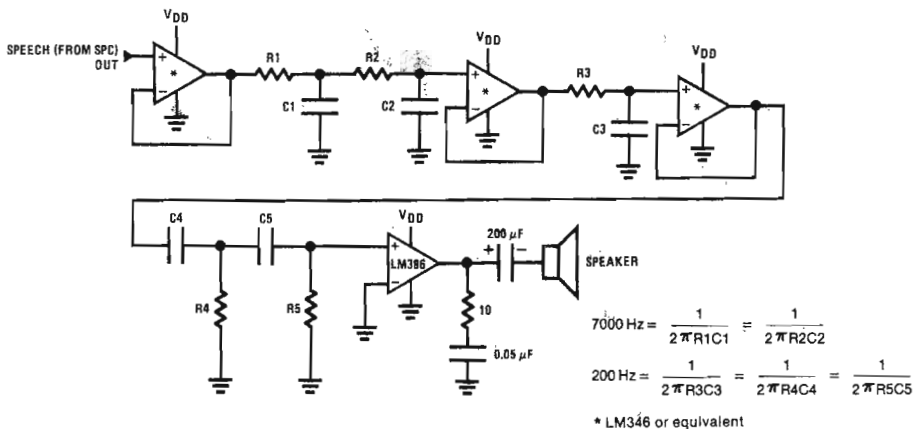
Note 3: This is optional filtering that can be eliminated for simpler systems. The acceptable range for this cutoff frequency is 6000 Hz-8000 Hz.

Typical Applications (Continued)

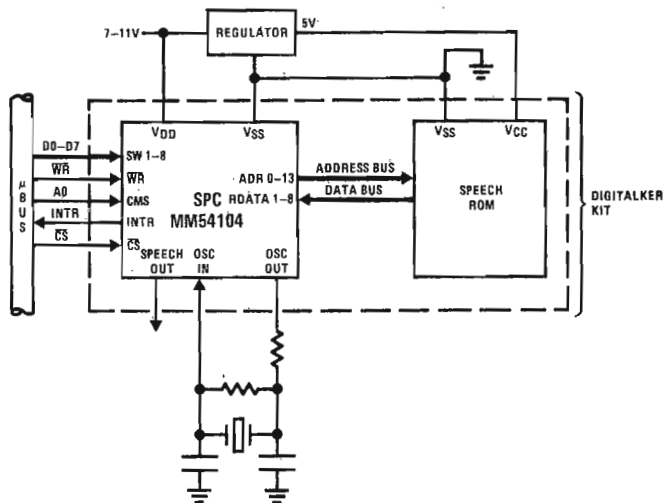
Minimum Filter Circuit



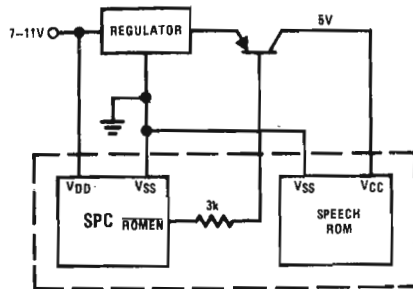
Filter Circuit to Produce Maximum Frequency Response



DIGITALTALKER™ System Utilizing MICROBUS™ Interface

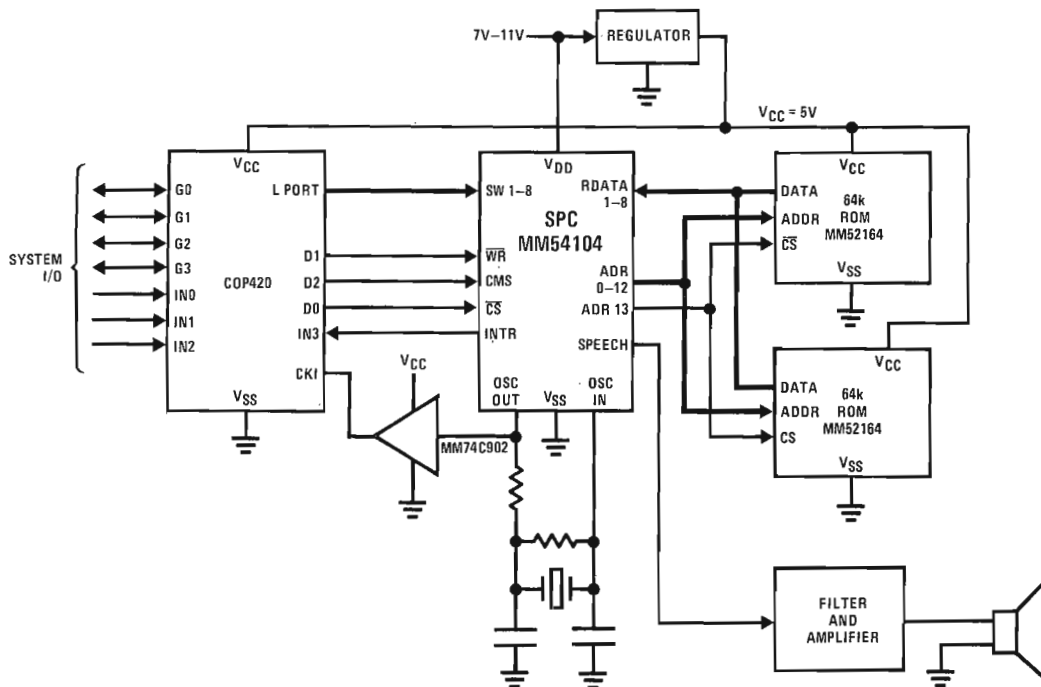


Low Power Configuration Using Static ROM

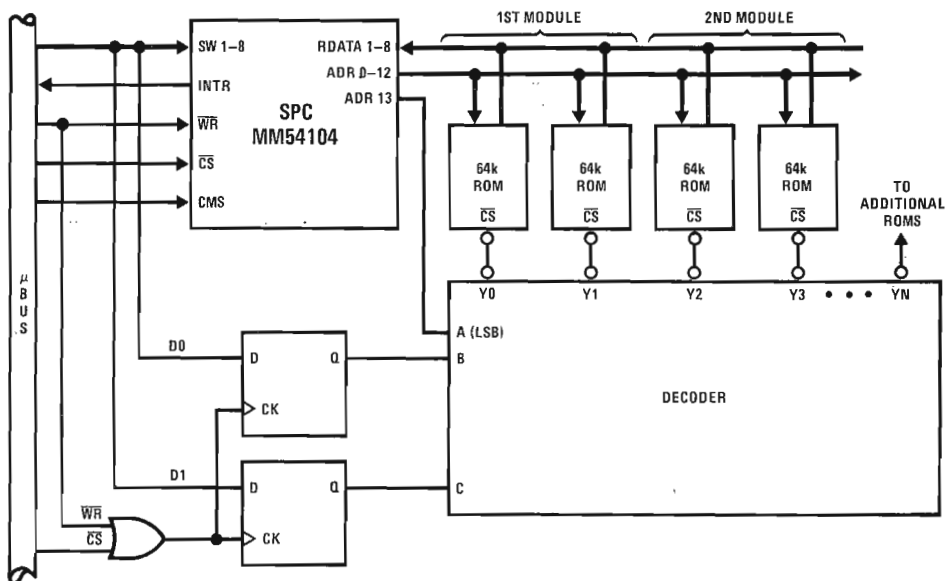


Typical Applications (Continued)

DIGITALKER™ System Using COP420 Interface



Speech ROM Expansion for Requirements Greater Than 128k Bits



Circuit for Evaluation of Custom Vocabulary EPROM Prototype Set

National Semiconductor
Linear Brief 54
Fred Wickersham
June 1981



EPROM PROTOTYPE

In the process of developing a product with a "custom" generated vocabulary, it may be necessary to develop special circuitry for listening to and evaluating your prototype synthesized vocabulary prior to committing to read only memory (ROM) production.

The prototype set will normally be supplied by National Semiconductor in the form of 2716 EPROM (Intel pinout) sets. The SPC (speech processor chip, part number MM54104) communication with EPROM sets *does* require some external hardware considerations which may not be necessary in the final ROM application, especially in multiple EPROM-equivalent ROM situations. (For example, four 16k-bit 2716 EPROMs equal one 64k-bit ROM.)

Shown on next page is a recommended circuit which shows proper interface between 2716 EPROMs and the SPC. The circuit covers vocabularies from the minimum system of one 2716 (16k bits) to larger vocabularies of eight 2716s (128k bits). It is also true, that in an application requiring only one 2716, the MM74LS138 decoder device can be eliminated by connecting pin 20 of the 2716 to V_{SS} (ground). The remaining unused pins 36, 37 and 38 of the SPC can be left unconnected in this case.

UNUSED SPC INPUTS

In any DIGITALKERTM design, an applications suggestion is in the area of unused input pins of the SPC. It should be understood that the number of different expressions and coincident addresses, as designated by the custom vocabulary, determines how many of the SW word address pins (pins 8-15) on the SPC are utilized. Vocabularies of less than 128 addresses will not use SW 8; vocabularies of less than 64 addresses will not use SW 7 or SW 8, and etc. These unused SW pins must be tied to V_{SS} (ground) to simplify the application. In fact, any unused input to the SPC must be tied to V_{SS} .

FILTERING

Use of the DIGITALKERTM is quite straightforward, however, a point on application that must be covered in this brief concerns the frequency response of the output speech. The ultimate quality of the DIGITALKERTM will strongly depend upon the filter, amplifier and speaker choices made

by the user. For that reason, it is important to understand the output characteristics of the device.

Because the synthesized speech data is derived from a differentiated and sampled input signal, it is necessary to pass the output waveform of the MM54104 through a low-pass filter with a cutoff frequency of approximately 200 Hz and an attenuation characteristic of 20 dB/decade. This compensates for the high frequency pre-emphasis used in the synthesis technique. If the system of interest has a natural rolloff near 200 Hz, this low-pass filter can be eliminated. The important item is that the entire audio system should have a cutoff frequency of approximately 200 Hz. The placement of the cutoff frequency may be adjusted for the particular type of voice being synthesized. A low pitched man's voice might sound better with a 100 Hz cutoff point while women's and children's voices may show improvements with a 300 Hz cutoff.

As an example of how the overall frequency response of a particular application can minimize the need for extra filtering, consider the DIGITALKERTM as a voice announcement circuit in a telephone system.

In this case, the telephone network provides a natural attenuation to high frequencies that balances the SPC high frequency pre-emphasis. As a result, the low-pass filter previously mentioned can be eliminated. However, because signal frequencies above 3 kHz must be attenuated before they are allowed to pass into the telephone network, a cutoff filter of 3400 Hz may be required in place of the previously mentioned 200 Hz low-pass filter. A good filter for this application is the National Semiconductor AF133 active filter.

In addition to the 200 Hz to 3400 Hz low-pass filter, an extra stage of low-pass filtering can be used for frequencies above 7 kHz. This filter is optional and is normally only used to further reduce sampling noise. Most systems can omit this filter, especially if the overall system bandwidth is not very wide. A second optional filter can be included to limit the overall low frequency response of the system. This high-pass filter would normally cutoff below 200 Hz (adjusted to match the 200 Hz low-pass if provided). This high-pass filter limits low frequency noise, and can usually be omitted if system characteristics do not require this function.

Speech Synthesis

National Semiconductor
Application Note 252
Jim Smith
Dave Weinrich



INTRODUCTION

Electronic speech circuits offer a new dimension of sophistication to many modern machines. As annunciators in trains, elevators, office buildings, autos, airplanes, terminals, toys and games, etc., electronic speech circuits provide a more direct and natural announcement than bells, buzzers or lights. With electronic voice signals, complex directions can be clearly given in any language and with a minimum of effort.

In the past, electronic announcement systems required elaborate tape mechanisms. These systems were expensive and troublesome, so their use was limited to the small number of applications that required speech announcements (e.g., telephone announcement systems). The first all-electronic systems used analog to digital conversion techniques to convert actual voice into digital signals. These digital speech signals were then stored as PCM or delta modulation signals in semiconductor memories. The major problem with this arrangement was the massive memory required for a moderate amount of announcement time. One second of digital speech, in this configuration, required from 16k to 100k bits of memory.

The latest solution to electronic speech is known as speech synthesis. This technique provides a dramatic reduction in the memory required for one second of speech. Memory requirements range from 400 bits to 2000 bits per second depending on the desired speech attributes and overall quality. The synthesizer technique takes advantage of the fact that speech signals are highly redundant and predictable. By coding only the slowly varying coefficients of speech or by dramatic compression of digitized speech, significant bandwidth reductions in the digitized signal can be realized. These techniques, coupled with LSI semiconductor technology, make true voice synthesis practical.

The National Semiconductor speech processor chip (SPC) provides the complete speech synthesis reproduction circuitry needed to generate high quality and natural speech (male, female or a child's voice), electronic tones or music. A complete chip set is called the DIGITALKERTM. It consists of the speech processor chip and a speech ROM. The applications for this chip set are endless, but to name a few:

- Voice interactive computer terminals
- Automotive, nautical and aeronautical instrumentation annunciators
- Voice-back units for banking, weather and time announcements, answering machines, etc.
- Elevators, trains, subway systems, etc.
- Consumer appliances, toys and games
- Warning systems for fire and police emergency

All of these applications benefit from the lower overall cost, high reliability, excellent performance and fast control response afforded by the National Semiconductor DIGITALKERTM system. The remainder of this note will be

devoted to a description of the MM54104 SPC, the technique used to synthesize speech and finally a review of the major DIGITALKERTM applications.

SPEECH SYNTHESIS

The basic phonological element of speech is the phoneme. The phoneme represents a simple sound that by itself cannot distinguish different words. Phonemes, together with speaker inflection, volume, emphasis, etc. are the fundamental building blocks of speech. The overall quality of any speech synthesizer, therefore, is directly controlled by its ability to faithfully reproduce all of the necessary speech attributes and not just phoneme reproduction.

The common American English language consists of approximately 38 to 40 phonemes—14-16 vowel sounds and 24 consonant sounds. Each phoneme is generated with either a voiced sound, as in "eye", or an unvoiced sound like the sh in "shy". This difference between a voiced and unvoiced sound is very important because the unvoiced sounds are generally fewer in number and less dependent upon the physiological characteristics of the speaker. A speech synthesizer, it turns out, can exploit this important difference. Finally, normal speech rates are approximately 10 to 15 phonemes per second (including silence intervals). Since 38-40 phonemes can be coded using 6 bits, the normal bit rate for phoneme reproduction is approximately 60 to 90 bits per second. This bit rate, however, contains only phoneme information which is only one of the many important speech attributes.

Since phoneme reproduction is a basic element in any speech synthesizer, an understanding of phoneme construction would be useful. Speech synthesis models use two driving functions, an impulse source for voiced sounds and a noise source (hiss noise) for unvoiced sounds. Each of these driving signals are filtered into specific frequency bands or formants by time-varying filters. The net result, for any particular set of valid filter coefficients, is a formant sound. In the human vocal tract system, the driving function consists of the lungs as the energy source, and the vocal cords for generating a voiced sound. The driving function for an unvoiced sound relies on the noise generated as air rushes through the vocal chambers and not on vocal cord vibrations. The formants are then generated by the resonant chambers of the throat, mouth and nasal cavities. By controlling the physical nature of these chambers with mouth position, tongue position and throat orifice size, a speaker can control the formants to generate a phoneme. It should be noted, however, that formants are identified by distinctive frequency bands. The unvoiced sounds do not generate these distinctive bands and therefore do not necessarily require the "normal" formants for a faithful reproduction. These sounds are characterized by a noise or hiss with very little resonance. This unvoiced resonance is normally identified as a fricative formant (e.g., the "sh" sound) and is characterized by an unusually large content of high frequencies.

A formant-based speech synthesizer, as described previously, would normally use at least three formant filters for voiced sounds and one formant filter for fricative sounds. An additional resonance, called nasal resonance, may be included but no dynamic formant function is usually associated with the nasal resonator. For the synthesis of a normal English vowel using a male voice, the three basic formants would fall into the approximate frequency bands of 200 Hz to 800 Hz, 900 Hz to 2300 Hz and 2400 Hz to 3000 Hz. The fricative formant is typically a pulse of high frequency noise in the band from approximately 2500 Hz to 8000 Hz, with the higher frequency fricatives like "th" usually much lower in relative amplitude when compared to the "sh" fricative sound.

The basic formant synthesizer requires formant filter coefficient data, amplitude control data and driving function control data. This minimum system could synthesize speech, but would not control inflection or emphasis. Its quality, therefore, can be very disappointing. Normal memory requirements for a minimal system are approximately 400 bits for one second of speech.

A second approach to speech synthesis does not automatically break speech into its minimum phonological elements. Instead, the speech waveform is sampled, digitized and compressed by the elimination of symmetrical redundancy and silent intervals, the use of adaptive delta modulation, and the adjustment of phase information in the digitized speech. In this way, speech elements can be synthesized as phonemes, phoneme groups, words or even whole phrases. Also, the attributes of the original speaker can be maintained if the synthesized elements are not broken down incorrectly (i.e., inflection can modify the sound of a phoneme if it occurs at the end of a word or phrase rather than at the beginning).

In a speech compression system, unvoiced sounds can be standardized. During the compression algorithm, the voiced and unvoiced sounds are separated and the voiced sounds are compressed. Unvoiced sounds, however, are compared to the available sounds and synthesized by substitution. This approach is successful because unvoiced sounds have very few speaker defined characteristics. As a result, a relatively small set of unvoiced sounds can be used repeatedly.

This speech compression technique offers excellent quality at a low data rate. The synthesis of a male voice, using English, will usually require an average of 1000 bits per word. Because the technique can be applied to any voice frequency signal, it is also capable of synthesizing women's and children's voices, music and tones. This flexibility, plus the realistic quality of the synthesized speech, make this technique very attractive.

THE NSC SPEECH PROCESSOR CHIP (SPC)

The National Semiconductor speech synthesis system consists of the SPC device plus the speech memory (ROM) required to assemble a complete DIGITALKER™ kit. To this kit a customer must add a clock input signal or the necessary oscillator components, an audio filter and amplifier and the control circuit function. This would represent the minimum configuration shown in Figure 1. The maximum amount of directly addressable speech memory accessible by the SPC is 128k bits, but external page addressing by the control circuit function can increase this ROM field as required.

The SPC utilizes the speech compression synthesis technique. As mentioned earlier, this technique reduces the amount of memory needed to store electronic speech by removing the excess or redundant data from the

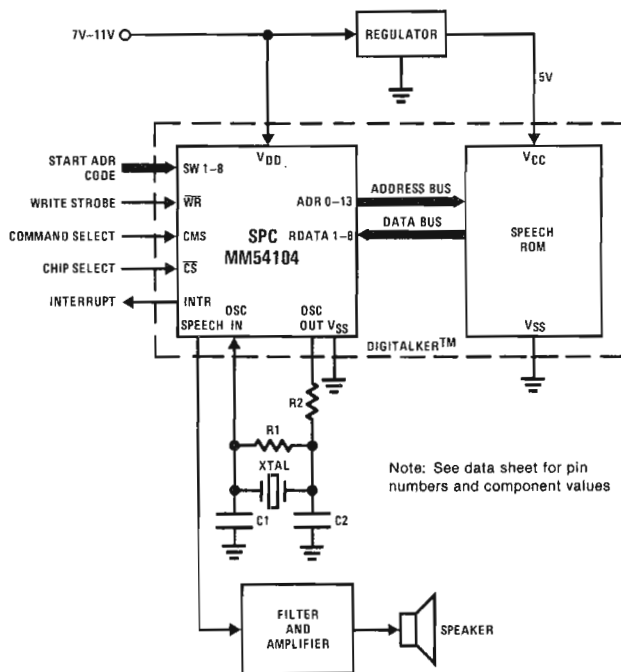


FIGURE 1. DIGITALKER™ Minimum Configuration

speech signal. The four main techniques to perform that task are:

1. Elimination of redundant pitch periods
2. Adaptive delta modulation coding to minimize bandwidth and memory requirements
3. Phase angle adjustments to create mirror image symmetry
4. Replacing the low level portion of a pitch period with silence (half-period zeroing)

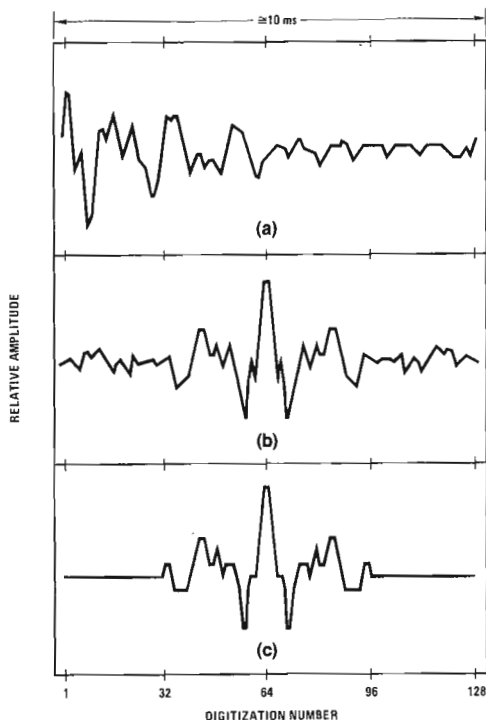
National Semiconductor uses an elaborate computer program to analyze a high fidelity tape recording and generate a ROM pattern that will faithfully synthesize the original voice message.

Figure 2 contains a block diagram of the MM54104 SPC. The eight-bit start address bus allows up to 256 separately defined sounds or expressions to be stored in the speech ROM. The control interface to the start address port can take the form of decoding logic, a MICROBUS™ port or mechanical switches.

When the $\overline{WR}$ goes high, the start address code is loaded into the control word address register. The SPC uses this control address to fetch the control word from ROM for the first block of speech data. The control word contains waveform information, repeat information and the address of the speech data. This address is loaded into the phoneme address register and is used to fetch the speech data used to recreate the speech waveform. Before the synthesis takes place, the waveform data must be decoded to provide information such as male or female, voiced or unvoiced, half-period zeroed or not half-period zeroed and silence.

The unsynthesized waveform for a typical voiced pitch period might look like the signal shown in Figure 3a. In the process of converting this signal to a synthetic form, several operations are performed. First, the phase delay of the signal can be adjusted to create a symmetrical waveform about the center of the pitch period as shown in Figure 3b. The next step will replace the low level beginning and ending quarters of the waveform with silence (Figure 3c). The result is a compression factor of 4 to 1 on

the original voice data. Now, delta modulation is applied and the results are shown in Figure 3c. Synthesis of the waveform starts with a period of silence (no speech data required), a quarter period of adaptive delta modulation-generated speech followed by the same speech data



(a) Original Speech Waveform
(b) Phase Angle Adjusted to Create Mirror Symmetry
(c) Half-Period Zeroed and Delta Modulated

FIGURE 3. SPC Waveforms (After Mozer [2])

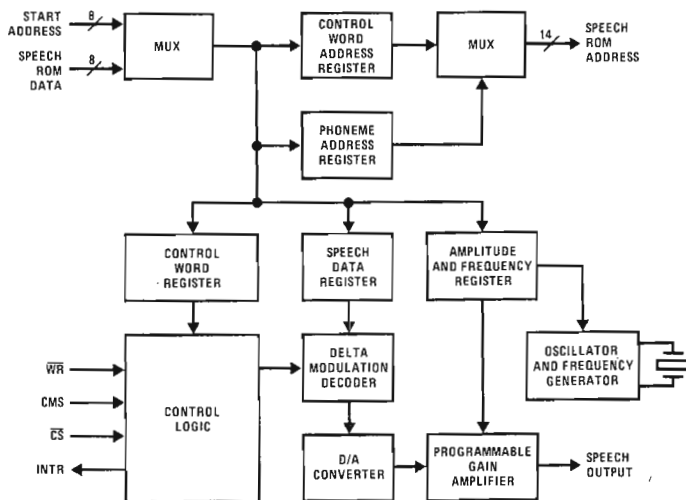


FIGURE 2. MM54104 Block Diagram

MICROBUS™ is a trademark of National Semiconductor Corp.

fetched in reverse. Finally, the SPC will finish the last quarter cycle of the speech block period with silence. This phase modified speech data sounds the same as the original speech.

At the end of a waveform or speech block, the SPC makes a decision about repeating the sequence. Each waveform of a typical voiced signal may be repeated an average of 3 to 4 times. The typical unvoiced waveform may be repeated approximately 7 to 8 times. Once the proper number of repeats has been generated, the SPC will begin a new speech block sequence. This operation continues until the SPC has executed all control words associated with the original eight-bit start address code.

SPC speech signals are stored as adaptive delta modulation data. This encoding technique exploits the relatively predictable and slowly changing characteristic of voiced speech. Because of the small differential between successive speech samples, a delta value rather than an absolute value can be used to determine the actual speech signal. Addition of the delta value to previously accumulated values will result in a new output waveform signal level. An adaptive technique is used so that the delta step size can change in response to slope variations. This technique uses multiple delta modulation step sizes to obtain a more accurate resolution and yet, the required amount of stored data remains lower than the information required for a more conventional encoding scheme.

The internal SPC clock is derived from a programmable frequency generator. Variations in the frequency of this clock, through the control word, allow the SPC to add a rising and falling pitch to speech sounds and syllables. This derived pitch variation adds a natural inflection to the synthetic speech.

Just as pitch variations are used to increase realism, so must the SPC use gain variations. Both techniques are controlled by data stored at the beginning of a speech block and the programmable oscillator and output amplifier circuit blocks of the SPC.

Use of the DIGITALKER™ is quite straightforward and will be outlined in the next section. However, a point on application that must be covered in this note concerns the frequency response of the output speech. The ultimate quality of the DIGITALKER™ will strongly depend upon the filter, amplifier and speaker choices made by the user. For that reason, it is important to understand the output characteristics of the device.

Because the synthesized speech data is derived from a differentiated and sampled input signal, it is necessary to pass the output waveform of the MM54104 through a low-pass filter with a cutoff frequency of approximately 200 Hz and an attenuation characteristic of 20 dB/decade. This compensates for the high frequency pre-emphasis used in the synthesis technique. If the system of interest has a natural rolloff near 200 Hz, this low-pass filter can be eliminated. The important item is that the entire audio system should have a cutoff frequency of approximately 200 Hz. The placement of the cutoff frequency may be adjusted for the particular type of voice being synthesized. A low pitched man's voice might sound better with a 100 Hz cutoff point while women's and children's voices may show improvements with a 300 Hz cutoff. Figure 4a shows a filter and amplifier circuit for this minimum frequency response characteristic.

As an example of how the overall frequency response of a particular application can minimize the need for extra

filtering, consider the DIGITALKER™ as a voice announcement circuit in a telephone system.

In this case, the telephone network provides a natural attenuation to high frequencies that balances the SPC high frequency pre-emphasis. As a result, the low-pass filter previously mentioned can be eliminated. However, because signal frequencies above 3 kHz must be attenuated before they are allowed to pass into the telephone network, a cutoff filter of 3400 Hz may be required in place of the previously mentioned 200 Hz low-pass filter. A good filter for this application is the National Semiconductor AF133 active filter.

In addition to the 200 Hz to 3400 Hz low-pass filter, an extra stage of filtering can be used for frequencies above 7 kHz. This filter is optional and is normally only used to further reduce sampling noise. Most systems can omit this filter, especially if the overall system bandwidth is not very wide. A second optional filter can be included to limit the overall low frequency response of the system. This high-pass filter would normally cutoff below 200 Hz (adjusted to match the 200 Hz low-pass if provided). This high-pass filter limits low frequency noise, and can usually be omitted if system characteristics do not require this function. A circuit having the full frequency response characteristic is shown in Figure 4b. Figure 5 shows the recommended overall speech synthesis system frequency response.

APPLICATIONS

While the variety of synthetic speech applications are numerous, the actual implementation in any single application is usually limited to one of the following three techniques:

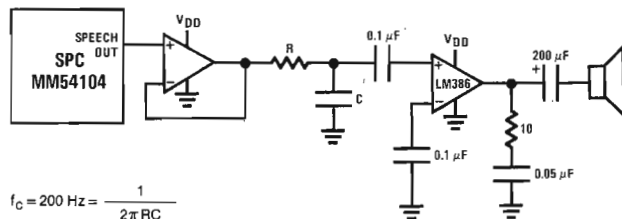
- (a) Single channel, hardware control logic
- (b) Single channel, software control logic
- (c) Multichannel, hardware or software control logic

Each of these circuit approaches for the SPC will be discussed in this section. Particular emphasis will be placed on items (b) and (c), however, because of the broad application possibilities for these two techniques.

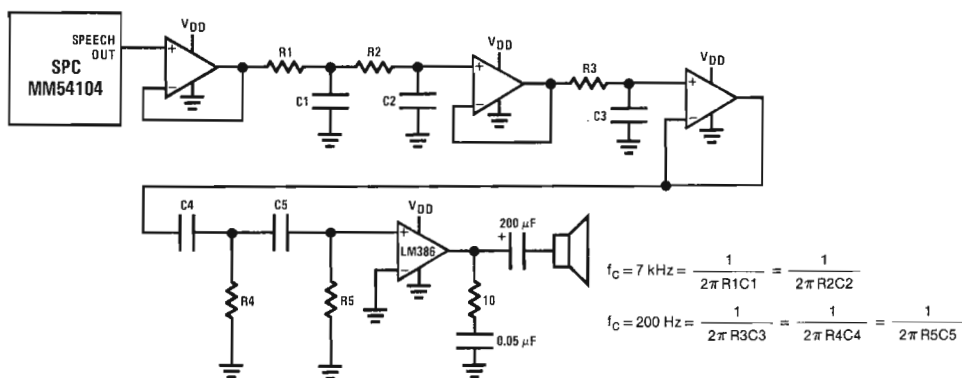
Certain applications require a relatively small number of sentences or announcements with very little similarity between the different sentences. An example of this application might be a talking elevator controller where the messages are brief and non-redundant (e.g., "going up, first floor, second floor", etc.). In this application, certain words are used repeatedly but the number of messages is limited and the length of each message is short. This application and others just like it, do not require the assembly of short phrases into complete sentences, nor do they require a dynamic message structure as would be required with an automatic bank teller (e.g., "your change is ten dollars") where a monetary amount may change from message to message. This fixed message application, therefore, may only require the minimum control circuit as shown in Figure 6.

In Figure 6, the SPC receives a separate coded input for each complete sentence or message that is synthesized. This input code is received by the SPC through the SW 1-8 port.

The circuit shown in Figure 6 uses a mechanical switch group to interface the SPC while the Figure 7 circuit uses a hardware logic controller to input the coded message control data.



(a) Minimum Low-Pass Filter/Amplifier



(b) Maximum Filter Response Configuration

FIGURE 4. SPC Filter and Amplifier

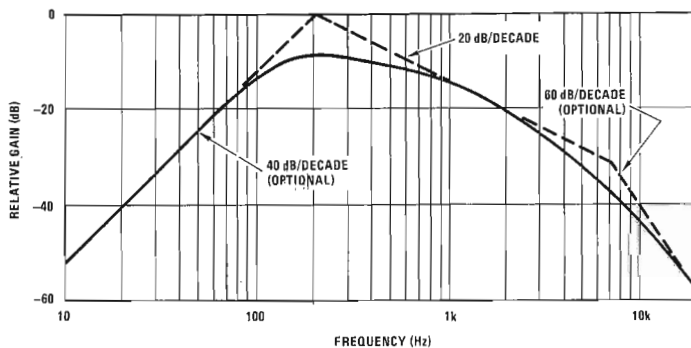


FIGURE 5. Recommended Frequency Response of Entire Audio System for MM54104 SPC

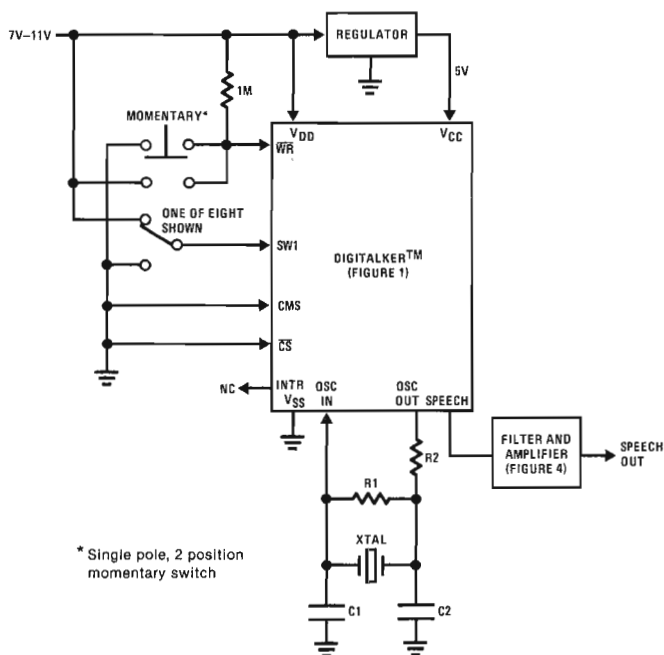


FIGURE 6. DIGITALKER™ with Switch Interface

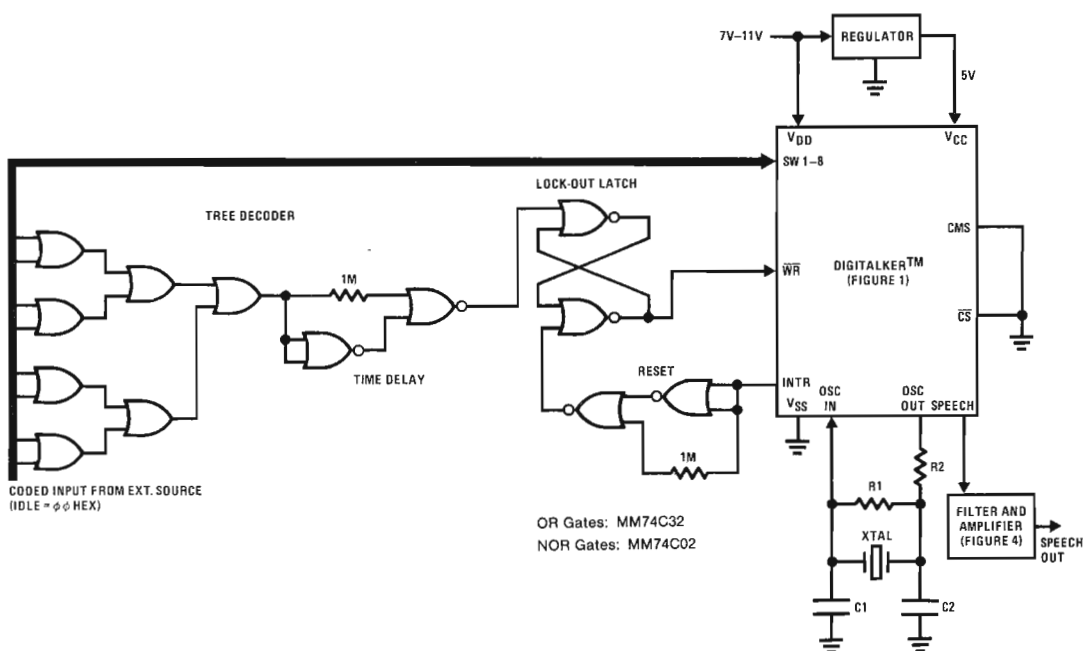


FIGURE 7. DIGITALKER™ with Logic Control Interface

After the proper message address is established on the SW 1-8 port, a momentary pulse must be applied to the WR line. If this signal is applied with a momentary action switch, as shown in *Figure 6*, then an external pull-up resistor should be used to pull the WR line up to logic high and complete the on-chip switch debounce circuitry. The suggested value of this resistance is one megohm. The WR input signal will latch the coded message address into the SPC on the rising edge of WR and initiate the synthetic speech message. Since each complete message uses a unique address code of the SW 1-8 port, no further control action is required after this point. The SPC will synthesize the requested message and return to the idle state. If a new input command signal is received, either during or after a message is synthesized, the SPC will immediately abort the current message and begin the new one. The circuit in *Figure 7* shows a lock-out circuit to prevent the aborting of a current message so that messages must be completed before a new message can be initiated.

In Figure 7, a message is initiated whenever a valid code word is applied to the eight-bit SW 1-8 port of the SPC. The valid code is detected by the combinational logic decoder and timed to insure all transitions have died. Once the valid code is timed, an S-R latch is set and a WR rising

edge is generated to start the SPC. This latch circuit also prevents retriggering of the SPC until after the present speech message is completed. Once the synthesized message has ended, the SPC will set the INTR line to the logic one state and a reset pulse will be generated to reset the lock-out latch. A new speech message can now be started by momentarily applying an idle address code followed by a valid code on the SW 1-8 input port.

The SPC will directly address up to 128k bits of speech memory. *Figure 8a* shows a typical speech ROM configuration of 128k using two 64k ROMs. The types of ROMs used have mask programmable chip selects, therefore, no extra decode logic is required for memory requirements of less than 128k. Although this memory size is usually sufficient for most applications, certain systems may require added speech ROM addressing. The circuit in *Figure 8b* shows how the speech ROM of an SPC kit can be expanded in 128k bit pages or modules. Each page is arranged to contain a complete portion of the entire speech library for a particular system. Each single speech data block, as addressed via the start address port of the SPC, must be contained within one ROM page. No page boundaries can be crossed during the synthesis of a speech expression.

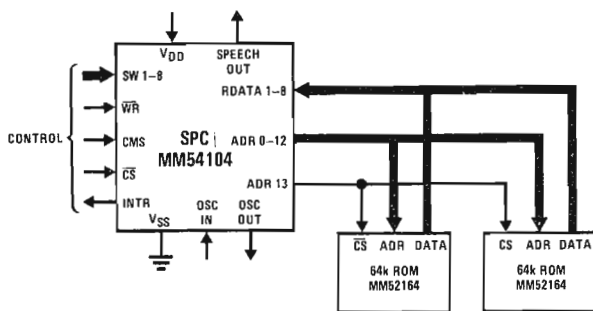


FIGURE 8a. Typical Speech ROM Configuration

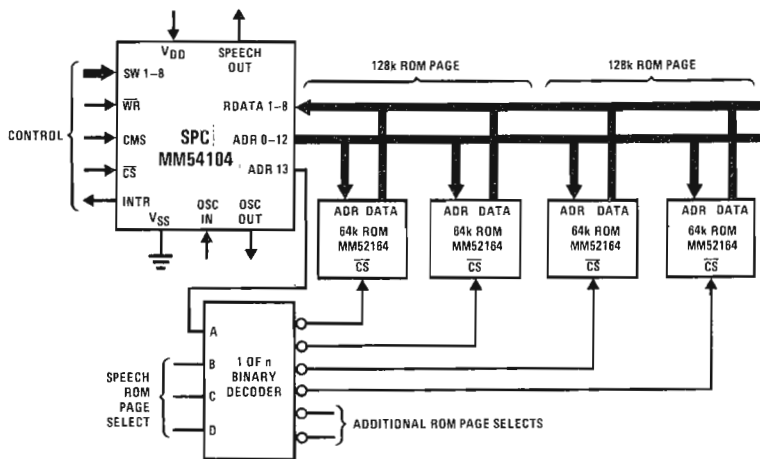


FIGURE 8b. Speech ROM Expansion Configuration

While the simple control schemes discussed so far can be used in many applications, a far more important group of applications will take advantage of the SPC's ability to construct sentences from a group of words, sounds and phrases. This type of application uses an intelligent controller or a microprocessor to string together a group of synthesized phrases to form a complete sentence. The electronic bank teller, previously mentioned, is a good example of this application. The microprocessor controls the stringing of SPC code addresses and applies them, one at a time, to the SW 1-8 port of the SPC. Handshake timing between the microprocessor and the SPC is provided with the INTR line. This microprocessor interface arrangement is known as MICROBUS™ and the configuration is shown in Figure 9.

The use of a microprocessor controller expands the versatility of the SPC tremendously. Messages that are composed of numerical responses or fixed phrases in random sequence can be easily constructed from a library speech memory. In addition, various tones or warnings can be synthesized and added before, during, or after an announcement to identify the urgency of each message. For example, an automobile message may state that "oil pressure is low". Alone, that message may only mean that pressure has dropped but no immediate hazard exists. If, however, pressure has dropped below a critical value, the message

could be compounded to say "warning, oil pressure is low, pull over and stop the engine". In this latter case, phrases of high urgency are added to the initial message to increase its level of importance. Of course, the second message is not completely separate from the first but is, instead, an expansion of the first. This technique allows fewer input address codes to initiate a larger number of messages without assigning a separate address code for each message and for each of its derivatives. This would be particularly important to an electronic bank teller since a large number of monetary amounts must be synthesized for a relatively small number of finished sentences.

When preparing a speech ROM for an SPC that will synthesize whole sentences from groups of phrases, it is important to note the desired inflections. The SPC has the ability to synthesize all of the important speech attributes including pitch and gain variations, emphasis, inflection, etc. This leads to very high quality life-like synthetic speech if the stringing of phrases does not result in an artificial emphasis or inflection. It is important to choose phrases carefully and to record them with the attribute required for a realistic sentence string. The stringing of phonemes should be avoided whenever possible because the natural inflection is usually lost in such an arrangement.

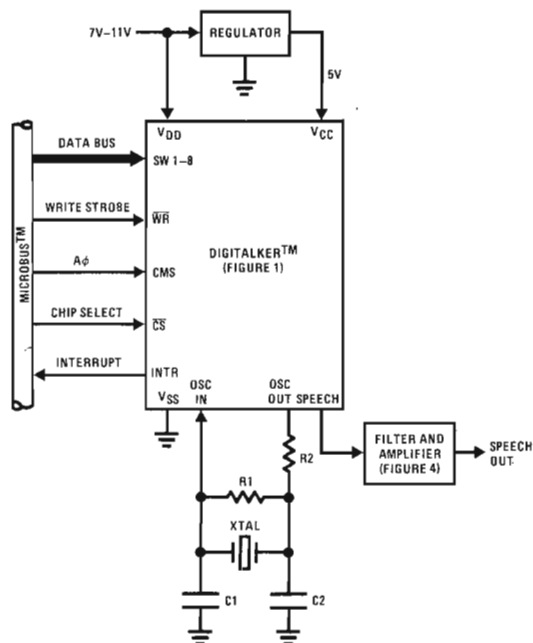


FIGURE 9. DIGITALTALKER™ with Microprocessor (MICROBUS™) Interface

A low cost intelligent controller for the SPC is one of the COP400 series of microcontrollers. Figure 10 shows one possible arrangement of an SPC system and a COP420. The COP provides all of the advantages associated with a MICROBUS™ interface at a relatively low cost. Because of its limited I/O structure, the COP's serial I/O port is expanded as required to obtain the desired number of input lines.

The final application technique to be covered is the multichannel configuration. The previous arrangements used an SPC and a dedicated set of speech ROMs to provide a single channel of synthetic speech. Appliances, autos, toys and games, terminals, etc. would probably use

a single channel SPC arrangement. But an entirely different group of products could take advantage of a multiple channel approach to reduce the ROM requirements. This group of products includes multiple elevator controllers, electronic bank tellers, multiple pupil learning centers, voice response telephone answering equipment, telephone switching system call announcement centers, etc. In this application, each channel would use a separate SPC and amplifier circuit, but several channels would share a common controller and speech library ROM. A typical configuration is shown in Figure 11.

The library ROM of Figure 11 is shared over eight SPC channels. Each SPC channel is scanned once in 16 μ s as

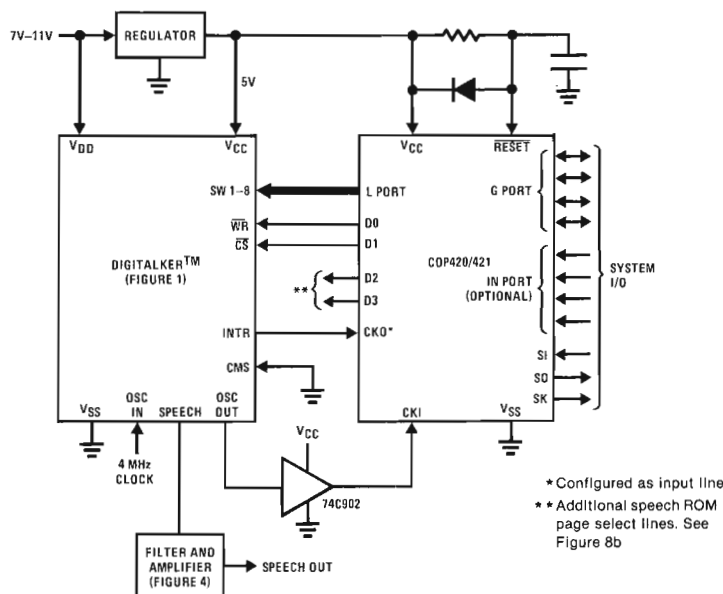


FIGURE 10. DIGITALK™ with COP420/COP421 Interface

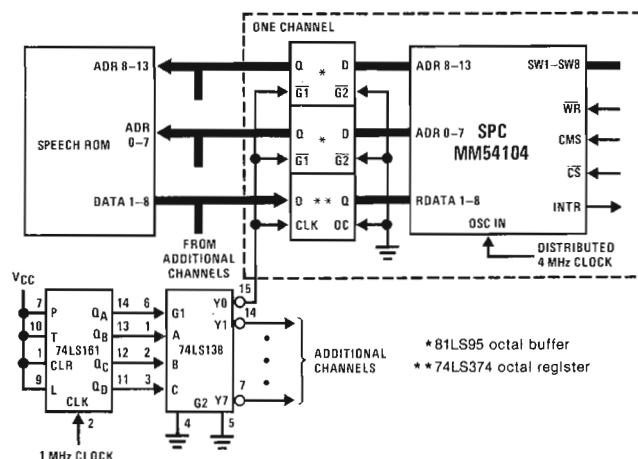


FIGURE 11. Multichannel Speech Synthesizer

shown in Figure 12. During each channel period of $2\mu\text{s}$, an SPC output address is presented to the ROM address input port via a pair of octal TRI-STATE[®] bus drivers. After one μs , the data from the ROM is clocked into the channel's octal data latch, the output of which is connected to the SPC ROM data input port. The remaining $1\mu\text{s}$ of each channel cycle is provided for bus settling time.

When the speech library ROM is shared over many channels, the actual number of shared channels is controlled by the MM54104 SPC memory cycle timing. Because the channel scanning is asynchronous to the SPC cycle timing, it is necessary for each channel to be scanned at least once during the high interval of the ROMEN signal. As shown in Figure 13, this signal is high for at least $20\mu\text{s}$ of each memory fetch cycle. Thus, a scanning rate of one channel every $16\mu\text{s}$ will insure that each channel is scanned at least once while the ROMEN signal is high.

One final note is necessary about the configuration in Figure 11. Simple modifications to the counter and decoder circuitry would allow this circuit to handle sixteen channels. A four-line to sixteen-line decoder would replace the three to eight decoder and the clock would directly enable the decoder during the logic low clock period. All sixteen channels would be scanned every $16\mu\text{s}$ and the scan interval for each channel would be one μs —

one-half μs of memory access time and one-half μs of bus guard time.

The last multichannel circuit is shown in Figure 14. This scheme reduces the number of wires needed between the speech ROM and each SPC channel. By multiplexing address and data over the same parallel bus, fewer wires are needed. This approach is particularly attractive when each SPC channel is located on an individual circuit card. A telephone central office or PABX announcement system is a typical example of a channel per card arrangement. Figure 14 represents that type of system.

As shown in Figure 15, each channel of the unified bus approach is scanned for one μs . As many as sixteen channels, therefore, can be scanned during the ROMEN high cycle of any SPC. During each channel scan, the bus is gated to transmit the ROM address to latches on the ROM circuit board. The address is sent in two bytes. After a brief delay of one-half μs , the bus is gated to return the requested ROM data to the same SPC channel. This data is then latched on the SPC channel card. This scheme is very straightforward. It exchanges reduced interconnect wiring for additional logic circuits.

To minimize interconnect wiring when using a unified bus structure, the SPC control logic would probably be configured on a per channel basis. The COP microcontroller,

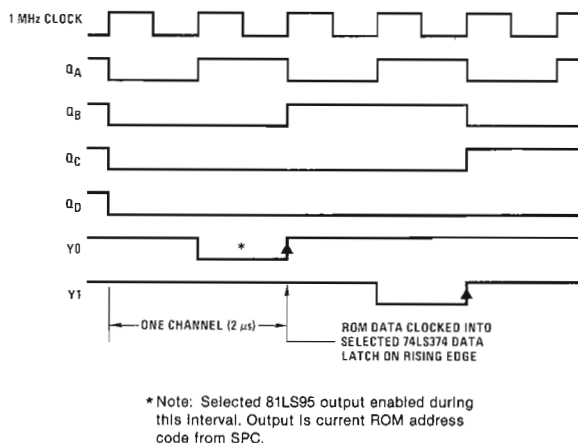


FIGURE 12. Multichannel Timing Diagram

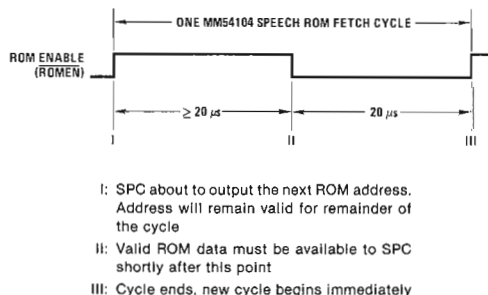


FIGURE 13. MM54104 SPC Speech Memory Cycle Timing

Once again, is a logical choice for this function. The COP controller initiates and assembles a group of fixed messages. Because of general similarities between the various messages, phrase strings are used to construct each finished message. Also, the circuit in Figure 14 allows one message to contain a non-fixed message—a telephone number. The COP controller reads a group of program switches or receives a down-loaded number from the switching system's central processor. It then inserts this number into the appropriate place during the syn-

thesis of the following typical message—"The number you are calling has been changed. The new number is 555-3434". The ROM library in this case contains the phrases required for message construction and the data needed to synthesize the name of each decimal digit. The library could also contain the names of the teen digit pairs, and the words "hundred" and "thousand". These would be used to synthesize the words "thirteen hundred" or "two thousand", etc.

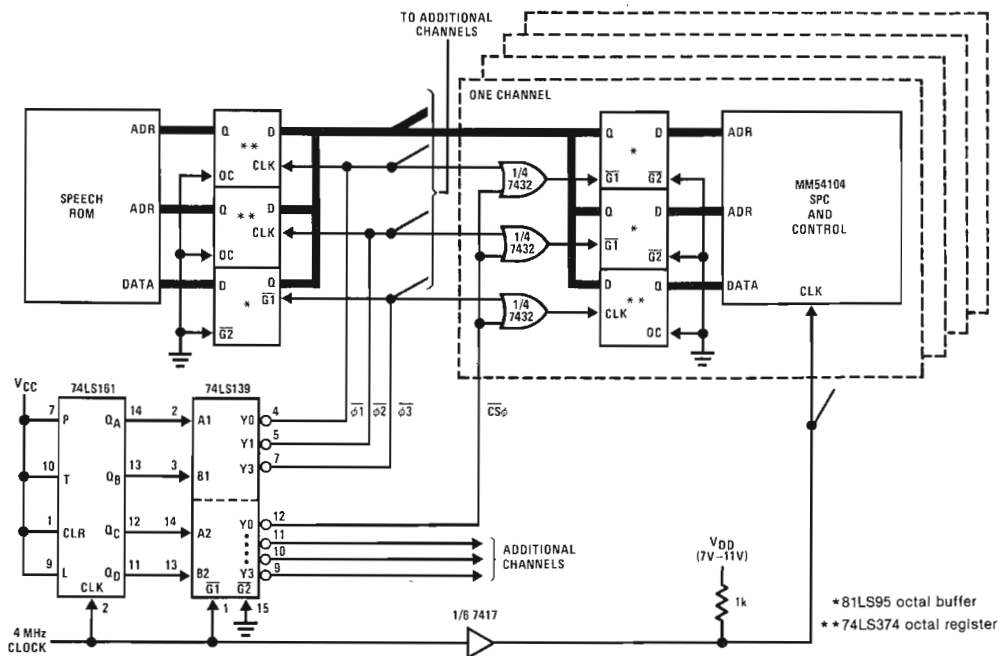


FIGURE 14. Multichannel Synthesizer with Unified Bus

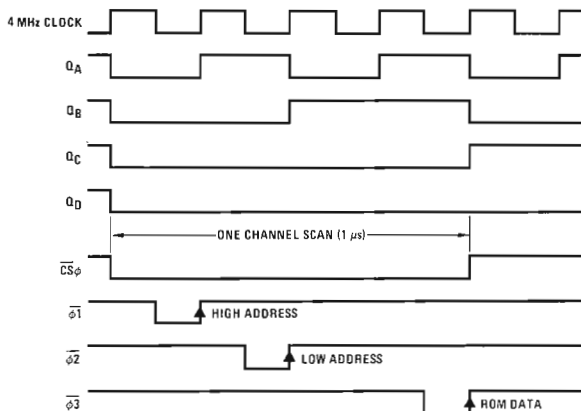


FIGURE 15. Multichannel (Unified Bus) Timing Diagram

SUMMARY

This application note describes some of the versatility and flexibility of the National Semiconductor DIGITALKER™ System. This system provides low cost speech and tone synthesis for products ranging from consumer items to computers. It provides a reliable alternative to mechanical systems (i.e., tape decks) without sacrificing voice quality. Also described in this note are a few of the most popular circuit arrangements possible with the DIGITALKER™. Of particular interest are the methods outlined to assemble whole messages from phrase groups and the schemes used for multichannel synthesizer systems. This latter application is particularly interesting because of the memory savings for the multichannel user.

REFERENCES

1. Morris, Dennis E. and Weinrich, David W., *A New Speech Synthesis Chip Set*, IEEE International Conference on Acoustics, Speech and Signal Processing, 1980.
2. Mozer, Forrest, *Method and Apparatus for Speech Synthesizing*, Pending US Patent.
3. Weinrich, David W., *A Speech Synthesis Chip Set Using Compression Techniques*, Electronics, April 10, 1980.

"The recipient of these products automatically receives a non-exclusive license under U.S. Patent Application 432,859 and any patent or patents issuing thereon to use such products, to assemble or otherwise incorporate them into further products which may be covered by said patent application, or any patent or patents issuing thereon, and to use, sell, or otherwise dispose of such products."

Protected by U.S. Pat. No. 4124125, F.M. Mozer licensee available.