DISTRIBUTORS

©2128-6827.5 Printed in U.S.A.

UNITED STATES

ALABAMA
Arrow (205) 830-1103
Hall-Mark (205) 839-8700

ARIZONA
(Phoenix Area)
Anthem (602) 968-6182
Hall-Mark (602) 243-6601

CALIFORNIA
(Northern)
Anthem (408) 946-8000
Arrow (408) 745-6600
Hall-Mark (408) 773-9990
Pacesetter (408) 734-5470

CALIFORNIA
(L. A. Area)
Anthem (714) 730-8000 (Tustin)
Anthem (213) 700-1000 (Chatsworth)
Arrow (714) 851-8961 (Orange County)
Arrow (213) 701-7500 (Chatsworth)
Pacesetter (714) 557-7131 (Santa Ana)

CALIFORNIA
(San Diego)
Anthem (714) 453-9008
Arrow (714) 565-4800
Hall-Mark (714) 268-1201

COLORADO
Arrow (303) 758-2100
Hall-Mark (303) 934-3111

CONNECTICUT
Arrow (203) 265-7741

FLORIDA
Arrow (305) 776-7790 (Ft. Lauderdale)
Arrow (305) 725-1480 (Palm Bay)
Hall-Mark (305) 971-9280 (Ft. Lauderdale)
Hall-Mark (305) 855-4020 (Orlando)
Hall-Mark (813) 576-8691 (St. Petersburg)

GEORGIA
Arrow (404) 449-8252
Hall-Mark (404) 448-5132

ILLINOIS
Advent (312) 297-6200
Arrow (312) 893-9420
Hall-Mark (312) 860-3800

INDIANA
Advent (317) 872-4910
Arrow (317) 243-9353

IOWA
Advent (319) 363-0221
Arrow (319) 395-7230

KANSAS
Hall-Mark (913) 888-4747

MARYLAND
Arrow (301) 247-5200
Hall-Mark (301) 796-9300

MASSACHUSETTS
Arrow (617) 933-8130
Lionex (617) 272-9400

MICHIGAN
Advent (313) 477-1650
Arrow (313) 971-8220

MINNESOTA
Arrow (612) 830-1800
Hall-Mark (612) 854-3223

MISSOURI
Arrow (314) 567-6888
Hall-Mark (314) 291-5350

NEW HAMPSHIRE
Arrow (603) 668-6968

NEW JERSEY
(Northern)
Arrow (201) 797-5800
Diplomat (201) 785-1830
Hall-Mark (201) 575-4415
Lionex (201) 227-7960

NEW MEXICO
Arrow (505) 243-4566

NEW YORK
(Long Island)
Arrow (516) 231-1000
Diplomat (516) 454-6400
Lionex (516) 273-1660

NEW YORK
(Upstate)
Arrow (716) 275-0300 (Rochester)
Arrow (315) 652-1000 (Syracuse)

NORTH CAROLINA
Arrow (919) 725-8711 (Winston-Salem)
Arrow (919) 876-3132 (Raleigh)
Hall-Mark (919) 832-4465 (Raleigh)

OHIO
Arrow (513) 435-5563 (Dayton)
Arrow (216) 248-3990 (Cleveland)
Hall-Mark (216) 473-2907 (Cleveland)
Hall-Mark (614) 891-4555 (Columbus)

OKLAHOMA
Hall-Mark (918) 665-3200

PHILADELPHIA AREA
Arrow (609) 235-1900
Hall-Mark (609) 424-0880
Pioneer (215) 674-4000

PITTSBURGH
Arrow (412) 856-7000

TEXAS
Arrow (512) 835-4180 (Austin)
Arrow (214) 386-7500 (Dallas)
Arrow (713) 491-4100 (Houston)
Hall-Mark (512) 258-8848 (Austin)
Hall-Mark (214) 341-1147 (Dallas)
Hall-Mark (713) 781-6100 (Houston)

UTAH
Arrow (801) 539-1135

WASHINGTON
Arrow (206) 643-4800

WISCONSIN
Arrow (414) 764-6600 (Milwaukee)
Arrow (608) 273-4977 (Madison)
Hall-Mark (414) 761-3000 (Milwaukee)

CANADA

Future (604) 438-5545 (Vancouver)
Future (514) 694-7710 (Montreal)
Future (416) 663-5563 (Toronto)
Future (613) 820-8313 (Ottawa)
Future (403) 259-6408 (Calgary)
R-A-E (604) 291-8866 (Vancouver)
R-A-E (403) 451-4001 (Edmonton)

AUSTRALIA

Daneva (03) 598-5622

Microelectronics Division/ WORLDWIDE SALES OFFICES

NORTH AMERICA
UNITED STATES:
MICROELECTRONICS DIVISION
NORTHEAST—600 West John Street
Hicksville, New York 11802
Tel: 516-733-3107, TWX: 510-221-1866

20th Century Plaza
Daniel Webster Highway
Merrimack, New Hampshire 03054
Tel: 603-424-3303, TWX: 710-366-0676

858 Welsh Road
Maple Glen, Pennsylvania 19002
Tel: 215-643-5326

SOUTHEAST—7901 4th Street. N., Suite 208
St. Petersburg, Florida 33702
Tel: 813-577-4024, TWX: 810-863-0398

1616 Forest Drive
Annapolis, Maryland 21403
Tel: 301-269-6250, TWX: 710-867-8566

4921C Professional Court
Raleigh, North Carolina 27609
Tel: 919-876-7380

SOUTH CENTRAL—5520 LBJ Frwy., Suite 330
Dallas, Texas 75240
Tel: 214-934-1654, TWX: 910-860-9259

CENTRAL—4524 S. Michigan Street
South Bend, Indiana 46614
Tel: 219-291-0585, TWX: 810-299-2518

5820 West 85th Street, Suite 102
Indianapolis, Indiana 46278
Tel: 317-872-7740, TWX: 810-341-3145

2355 S. Arlington Hts. Road, Suite 408
Arlington Heights, Illinois 60005
Tel: 312-981-0040, TWX: 910-687-0254

32969 Hamilton Court, Suite 210
Farmington Hills, Michigan 48018
Tel: 313-391-4070

SOUTHWEST—201 Standard Street
El Segundo, California 90245
Tel: 213-322-7745, TWX: 910-348-6296

NORTHWEST—3080 Olcott Street, Suite 230C
Santa Clara, California 95051
Tel: 408-496-0844, TWX: 910-379-0010

EUROPE
EUROPEAN SALES HEADQUARTERS:
GENERAL INSTRUMENT MICROELECTRONICS LTD.
Regency House, 1-4 Warwick Street
London W1R 5WB
Tel: 01-439-1895, Telex: 23272

NORTHERN EUROPEAN SALES OFFICE:
Regency House, 1-4 Warwick Street
London W1R 5WB
Tel: 01-439-1891, Telex: 23272

Sandhamnsgatan 67
S-115 28, Stockholm
Tel: 08-67 99 25, Telex: 17779

CENTRAL EUROPEAN SALES OFFICE:
GENERAL INSTRUMENT DEUTSCHLAND GmbH
(MOS Produktgruppe)
Nordendstrasse 3, 8000 Munchen 40
Tel: (089)27 24 049, Telex: 528054

6070 Langen Bei Frankfurt A Main
Wilhelm-Leuschner Platz 8, Postf. 1167
Tel: (6103) 23 051, Telex: 415000

SOUTHERN EUROPEAN SALES OFFICE:
5-7 Rue De L'Amiral Courbet
94160 Sainte Mande, Paris
Tel: 365 72 50, Telex: 213073

Piazza Novelli, 8
20129 Milano
Tel: 720-914, Telex: 843-320348

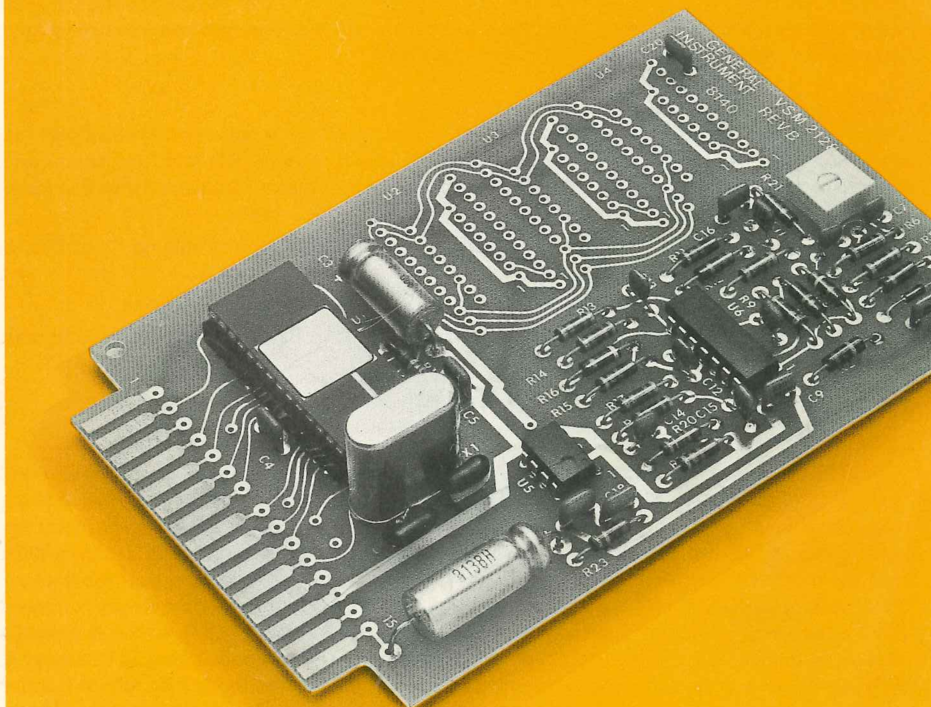
ASIA
HONG KONG:
GENERAL INSTRUMENT HONG KONG LTD.
139 Connaught Road Central, 3/F
San-Toi Building
Tel: (5) 434360, Telex: 84606

JAPAN:
GENERAL INSTRUMENT INTERNATIONAL CORP.
Fukide Bldg. 8th Floor, 1-13 Toranomon 4-Chome
Minato-ku, Tokyo 105
Tel: (03) 437-0281, Telex: 2423413

TAIWAN:
GENERAL INSTRUMENT MICROELECTRONICS TAIWAN
77 Pao Chiao Road, Hsin Tien
Taipei, Taiwan
Tel: (02) 914-6234, Telex: 785-3111

GENERAL INSTRUMENT

VSM2128 VOICE SYNTHESIS MODULE USER'S GUIDE



Orator™ VSM2128

GENERAL INSTRUMENT

Printed in U.S.A.

GENERAL INSTRUMENT

Voice Synthesis Module — VSM2128

FEATURES

- Complete Speech System
- Allophone Library
- Create Any Phrase
- Simple Digital Interface (TTL)
- 4.6 to 7.0 V Power Supply
- Audio Output: 200mW
- Operating Temperature: 0° to 70° C
- Dimensions: 3.25" x 5.0"

DESCRIPTION

The VSM2128 utilizes the General Instrument SP0256 Speech Processor. This device is a single chip, N-Channel, MOS/LSI circuit that is able, through stored program, to synthesize speech.

The module is easily interfaced with any digital system; ten TTL compatible signals are used to select the allophones. Once selected, the VSM2128 requires no support from the user circuit. It enunciates the allophone and signals when complete.

INTERFACE

The VSM2128 is interfaced using a 15-pin card edge connector (Amphenol 225-21521-401 (117)) or equivalent. The allophone is selected with an 8 bit address (S0-S7) and is strobed into the module by a low to high pulse on the $\overline{\text{STROBE}}$ line. The module will drive the $\overline{\text{BUSY}}$ line low and the $\overline{\text{LOAD}}$ line high. During this time, new data will not be accepted.

When the $\overline{\text{BUSY}}$ line is high this indicates that the module is inactive (not talking). The $\overline{\text{LOAD}}$ line is low whenever the input buffer is empty.

The module is initialized by applying a low-to-high transition to the $\overline{\text{RESET}}$ line.

The Audio is filtered by an eight-pole Butterworth Filter to insure the highest quality and is amplified to drive an 8Ω load with 200mW of power.

ALLOPHONE SPEECH SYNTHESIS

INTRODUCTION

The General Instrument allophone speech synthesis technique is easy to use, has a remarkably low bit rate, and allows the user to synthesize any English word by concatenating individual speech sounds. Each allophone requires a six bit address. Assuming that speech contains ten to twelve allophones per second, allophone synthesis would require addressing less than 100 bits per second. Previous techniques have involved synthesizing and storing entire words as units. The major disadvantage of this method is that, unless use of a very large memory is intended, vocabulary size is limited. For example, pulse code modulation (PCM), which is essentially digital recording, storage, and playback of speech waveforms, requires about 70 thousand data bits per second of speech. Another method, linear predictive coding (LPC), which predicts a speech sample from a weighted combination of previous samples, requires only one to two thousand bits per second of speech. Using this method, approximately 15-20 words can be stored in 16K bits of memory. While these methods require a large memory for a limited vocabulary, their main advantage is relatively high quality speech.

In contrast, allophone synthesis has the major advantage of providing an unlimited vocabulary, since the stored units are not words, but individual speech sounds (allophones). The user has merely to become familiar with the speech sounds of English (which are different from letters) and the allophone symbols used to represent them. Another use for allophone synthesis is in a text-to-speech system in which the user inputs a string of text (such as this). The advantage of such a system is that the user does not have to learn the allophone symbols of the two algorithms which are required: one which converts text to allophone symbols, and a second which converts those symbols to sounds. It is this second format that is being presented.

One disadvantage of allophone synthesis is that although completely intelligible, the speech quality is not as good as for PCM or LPC. The quality problem arises through concatenation of the allophones to form words. This will be discussed further in the sections to follow.

LANGUAGE

In order to successfully use a set of allophone sounds to synthesize words, there are a few preliminary points which should be made about speech and language. First, there is no one-to-one correspondence between written letters and speech sounds; second, speech sounds are not discrete units such as beads on a string; and last, speech sounds are acoustically different, dependent upon position within a word.

The first point compares to what a child encounters when learning to read. Each sound in a language may be represented by more than one letter and, conversely, each letter may represent more than one sound. (See the examples in Table 1). Because of these spelling irregularities, it is necessary to think in terms of **sounds**, not letters, when dealing with speech allophones.

The second point concerns segmentation of the speech signal. An adult who has learned to read usually thinks of the acoustic stream of speech as a string of discrete sounds which he calls by their letter names. In fact, speech is a continuously varying signal which cannot easily be broken into distinct sound-size units. For example, if an attempt is made to extract the **b** sound from the word **bat** by taking successively larger chunks of the acoustic signal from the beginning of the word, either a non-speech noise or the syllable **ba** is heard. In other words, there is no point at which the **b** sound can be heard in isolation.

The final point, and the most important for users of an allophone set, is that the acoustic signal of a speech sound may differ depending upon word positions. For example, the initial **p** in **pop** will be acoustically different from the **p** in **spy**, and may be different from the final **p** in **pop**. The ear will perceive the same acoustic signal differently depending upon which sounds precede or follow it. The word **cot** can be made to sound like **cod** by lengthening the duration of the **o** and, conversely, the word **cod** can be made to sound like **cot** by shortening the duration of the **o**.

PHONEMES OF ENGLISH

The sounds of a language are called phonemes, and each language has a set which is slightly different from that of other languages. Table 2 contains a chart of all the consonant phonemes of English, Table 3 all the vowel phonemes of English.

Consonants are produced by creating a constriction or occlusion in the vocal tract which produces an aperiodic sound source. If the vocal cords are vibrating at the same time, as in the case of the voiced fricatives VV, DH, ZZ, and ZH, (See Table 4) there are two sound sources: one which is aperiodic and one which is periodic.

Vowels are produced with a relatively open vocal tract and a periodic sound source (unless they are whispered) provided by the vibrating vocal cords. Vowels are classified according to whether the front or back of the tongue is high or low (See Table 3), whether they are long or short, and whether the lips are rounded or unrounded. In English, all rounded vowels are produced in or near the back of the mouth (UW, UH, OW, AO, OR, AW).

Speech sounds which have features in common behave in similar ways. For example, the voiceless stop consonants PP, TT, and KK (See Table 2) should be preceded by 50-80 msec of silence, and the voiced stop consonants BB, DD, and GG by 10-30 msec of silence.

ALLOPHONES

Phoneme is the name given to a group of similar sounds in a language. Recall that a phoneme is acoustically different depending upon word position. Each of these positional variants is an allophone of the same phoneme. An allophone, therefore, is the manifestation of a phoneme in the speech signal. It is for this reason that the inventory of English speech sounds is called an allophone set.

HOW TO USE THE ALLOPHONE SET

(See Table 5 for instructions on how to create all the sample words mentioned in this section.)

The allophone set (See Table 4) contains two or three versions of some phonemes. It may be necessary to use one allophone of a particular phoneme for word-or syllable-initial position and another for word-or syllable-final position. A detailed set of guidelines for using the allophones is given in Table 6. Note that these are suggestions, not rules.

For example, DD2 sounds good in initial position and DD1 sounds good in final position, as in "daughter" and "collide". One of the differences between the initial and final versions of a consonant may be that an initial version is longer. Therefore, to create an initial SS, two SSs can be used instead of the usual single SS at the end of a word or syllable, as in "sister". Note that this can be done with TH, and FF, and the inherently short vowels (to be discussed below), but with no other consonants. Experiment with some consonant clusters (strings of consonants such as **str, cl**) to discover which version works best in the cluster. For example KK1 sounds good before LL as in "clown", and KK2 sounds good before WW as in "square". One allophone of a particular phoneme may sound better before or after back vowels and another before or after front vowels. KK3 sounds good before UH and KK1 sounds good before IY, as in "cookie". Some sounds (PP, BB, TT, DD, KK, GG, CH, and JH) require a brief duration of silence before them. For most of these, the silence is included in the allophone, but more may be added as desired. Therefore, there are several pauses in the allophone set varying from 10-200 msec. To create the final sounds in the words "letter" and "little" use the allophones ER and EL. Think about how a word **sounds**, not how it is spelled. For example, the NG allophone obviously belongs at the end of the words "sing" and "long", but notice that the NG sound is represented by the letter N in "uncle". Some sounds may not be represented in words by any letters, such as the YY sound in "computer".

As mentioned earlier, there are some vowels which can be doubled to make longer versions for stressed syllables. These are the inherently short vowels IH, EH, AE, AX, AA, and UH. For example, in the word "extent" one EH is used in the first syllable, which is unstressed, and two EHs are used in the second syllable, which is stressed. Of the inherently long vowels there is one, UW, which has a long and short version. The short one, UW1, sounds good after YY in computer. The long version, UW2, sounds good in monosyllabic words like "two". Included in the vowel set is a group called R-colored vowels. These are vowel + R combinations. For example, the AR in "alarm" and the OR in "score". Of the R-colored vowels there is one, ER, which has a long and short version. The short version is good for polysyllabic words with final ER sounds like "letter", and the long version is good for monosyllabic words like "fir". One final suggestion when creating sentences is to add a pause of 30-50 msec between words and a pause of 100-200 msec between clauses.

Table 1 EXAMPLES OF SPELLING IRREGULARITIES

	Same sound represented by different letters	Different sounds represented by the same letter(s)
Vowels	m**ea**t f**ee**t P**e**te p**eo**ple p**e**nny	v**ei**n for**ei**gn d**ei**sm d**ei**cer g**ei**sha
Consonants	**sh**ip ten**s**ion pre**c**ious na**t**ion	al**though** **gh**astly cou**gh**

Table 2 CONSONANT PHONEMES OF ENGLISH**

		Labial[1]	Labio-Dental[2]	Inter-Dental[3]	Alveolar[4]	Palatal[5]	Velar[6]	Glottal[7]
Stops:	Voiceless Voiced	PP BB			TT DD		KK GG	
Fricatives:	Voiceless Voiced	WH	FF VV	TH DH	SS ZZ	SH ZH*		HH
Affricates:	Voiceless Voiced					CH JH		
Nasals:	Voiced	MM			NN		NG*	
Resonants:	Voiced	WW			RR, LL	YY		

*These do not occur in word-initial position in English.
**Examples of these phonemes in word context can be found in Table 5.

1. Upper and Lower Lips Touch or Approximate
2. Upper Teeth and Lower Lip Touch
3. Tongue Between Teeth
4. Tip of Tongue Touches or Approximates Alveolar Ridge (just behind upper teeth)
5. Body of Tongue Approximates Palate (roof of mouth)
6. Body of Tongue Touches Velum (posterior portion of roof of mouth)
7. Glottis (opening between vocal cords)

Table 3 VOWEL PHONEMES OF ENGLISH

	FRONT	CENTRAL	BACK
HIGH	YR IY IH*		UW# UH*#
MID	EY EH* XR	ER AX*	OW# OY#
LOW	AE*	AW# AY AR AA*	AO*# OR#

* SHORT VOWELS # ROUNDED VOWELS

Table 4 ALLOPHONES

< Silence >

PA1	PAUSE	10MS
PA2	PAUSE	30MS
PA3	PAUSE	50MS
PA4	PAUSE	100MS
PA5	PAUSE	200MS

< Short Vowels >

*/IH/	SiT	50MS
*/EH/	eND	50MS
*/AE/	HaT	80MS
*/UH/	BooK	70MS
*/AO/	auGHT	70MS
*/AX/	SuCCEED	50MS
*/AA/	HoT	60MS

< Long Vowels >

/IY/	See	170MS
/EY/	BeiGE	200MS
/AY/	SKy	170MS
/OY/	Boy	290MS
/UW1/	To	60MS
/UW2/	To	170MS
/OW/	Beau	170MS
/AW/	ouT	250MS
/EL/	SADDle	140MS

< R - Colored Vowels >

/ER1/	Fir	110MS
/ER2/	Fir	210MS
/OR/	STore	240MS
/AR/	ALarM	200MS
/YR/	CLear	250MS
/XR/	REPair	250MS

< Resonants >

/WW/	wOOL	150MS
/RR1/	rURAL	130MS
/RR2/	BrAIN	80MS
/LL/	lAKE	80MS
/YY1/	yES	90MS
/YY2/	yES	130MS

< Voiced Fricatives >

/VV/	vEST	130MS
/DH1/	thEY	140MS
/DH2/	thEY	180MS
/ZZ/	zOO	150MS
/ZH/	AzURE	130MS

< Voiceless Fricatives >

*/FF/	fOOD	110MS
*/TH/	thIN	130MS
*/SS/	VEsT	60MS
/SH/	shIP	200MS
/HH1/	hE	90MS
/HH2/	hOE	130MS
/WH/	whIG	150MS

< Voiced Stop Consonants >

/BB1/	bUSINESS	40MS
/BB2/	bUSINESS	60MS
/DD1/	COULd	50MS
/DD2/	dO	80MS
/GG1/	gUEST	80MS
/GG2/	gOT	80MS
/GG3/	WIg	120MS

< Voiceless Stop Consonants >

/PP/	pOW	150MS
/TT1/	PARt	80MS
/TT2/	tO	100MS
/KK1/	cAN'T	120MS
/KK2/	SkY	140MS
/KK3/	cOMB	80MS

< Affricates >

/CH/	chURCH	150MS
/JH/	DOdgE	100MS

< Nasals >

/MM/	mILK	180MS
/NN1/	THIn	170MS
/NN2/	nO	140MS
/NG/	AnCHOR	200MS

*These Allophones can be doubled.

Table 5 EXAMPLES OF WORDS MADE FROM ALLOPHONES

DD2-AO-TT2-ER1	"daughter"
KK3-AX-LL-AY-DD1	"collide"
SS-SS-IH-SS-TT2-ER1	"sister"
KK1-LL-AW-NN1	"clown"
KK3-UH-KK1-IY	"cookie"
LL-EH-TT2-ER	"letter"
LL-IH-TT2-EL	"little"
AX-NG-KK3-EL	"uncle"
KK1-AX-MM-PP1-YY1-UW1-TT2-ER	"computer"
EH-KK1-SS-TT2-EH-EH-NN1-TT2	"extent"
TT2-UW2	"two"
AX-LL-AR-MM	"alarm"
SS-KK3-OR	"score"
FF-ER2	"fir"

Table 6 GUIDELINES FOR USING THE ALLOPHONES

Silence	
PA1 (10ms)	before BB, DD, GG, and JH
PA2 (30ms)	before BB, DD, GG, and JH
PA3 (50ms)	before PP, TT, KK, and CH, and between words
PA4 (100ms)	between clauses and sentences
PA5 (200ms)	between clauses and sentences
Short Vowels	
*/IH/	sitting, stranded
*/EH/	extent, gentlemen
*/AE/	extract, acting
*/UH/	cookie, full
*/AO/	talking, song
*/AX/	lapel, instruct
*/AA/	pottery, cotton
Long Vowels	
/IY/	treat, people, penny
/EY/	great, statement, tray
/AY/	kite, sky, mighty
/OY/	noise, toy, voice
/UW1/	after clusters with YY: computer
/UW2	in monosyllabic words: two, food
/OW/	zone, close, snow
/AW/	sound, mouse, down
R-Colored Vowels	
/ER1/	letter, furniture, interrupt
/ER2/	monosyllables: bird, fern, burn
/OR/	fortune, adorn, store
/AR/	farm, alarm, garment
/YR/	hear, earring, irresponsible
/XR/	hair, declare, stare
Resonants	
/WW/	we, warrant, linguist
/RR1/	initial position: read, write, x-ray
/RR2/	initial clusters: brown, crane, grease
/LL/	like, hello, steel
/EL/	little, angle, gentlemen
/YY1/	clusters: cute, beauty, computer
/YY2/	initial position: yes, yarn, yo-yo
Voiced Fricatives	
/VV/	vest, prove, even
/DH1/	word-initial position: this, then, they
/DH2/	word-final and between vowels: bathe, bathing
/ZZ/	zoo, phase
/ZH/	beige, pleasure
Voiceless Fricatives	
*/FF/	These may be doubled for initial position and used singly in final position
*/TH/	
*/SS/	
/SH/	shirt, leash, nation
/HH1/	before front vowels: YR, IY, IH, EY, EH, XR, AE
/HH2/	before back vowels: UW, UH, OW, OY, AO, OR, AR
/WH/	white, whim, twenty

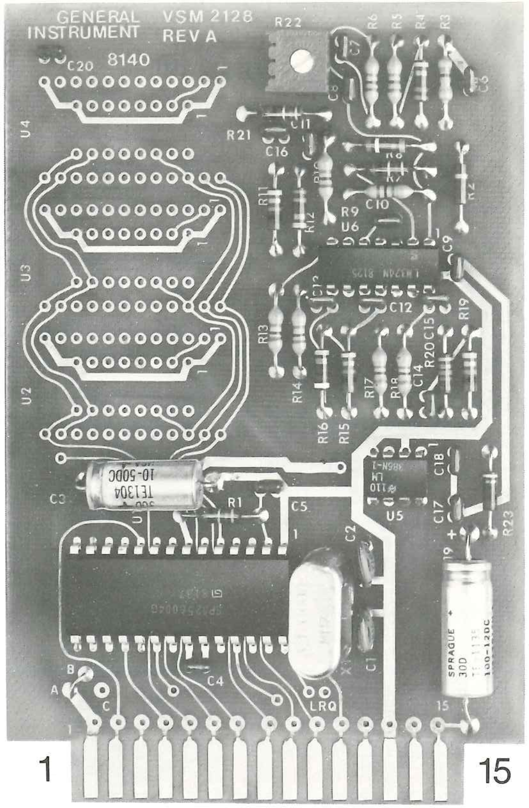
TABLE 6 (cont)

Voiced Stops	
/BB1/	final position: ri**b**; between vowels: fi**bb**er; in clusters; **bl**eed, **br**own
/BB2/	initial position before a vowel: **b**east
/DD1/	final position: play**ed**, en**d**
/DD2/	initial position: **d**own; clusters: **dr**ain
/GG1/	before high front vowels: YR, IY, IH, EY, EH, XR
/GG2/	before high back vowels: UW, UH, OW, OY, AX; and clusters: **gr**een, **gl**ue
/GG3/	before low vowels: AE, AW, AY, AR, AA, AO, OR, ER; in medial clusters: an**g**er; and final position: pe**g**
Voiceless Stops	
/PP/	**p**leasure, am**p**le, tri**p**
/TT1/	final clusters before SS: tes**ts**, i**ts**
/TT2/	all other positions: **t**est, s**t**reet
/KK1/	before front vowels: YR, IY, IH, EY, EH, XR, AY, AE, ER, AX; initial clusters: **c**ute, **cl**own, s**cr**eam
/KK2/	final position: spea**k**; final clusters: tas**k**
/KK3/	before back vowels: UW, UH, OW, OY, OR, AR, AO; initial clusters: **cr**ane, **qu**ick, **cl**own, s**cr**eam
Affricates	
/CH/	**ch**ur**ch**, feature
/JH/	**j**u**dg**e, in**j**ure
Nasal	
/MM/	**m**ilk, alar**m**, a**m**ple
/NN1/	before front and central vowels: YR, IY, IH, EY, EH, XR, AE, ER, AX, AW, AY, UW; final clusters: ear**n**
/NN2/	before back vowels: UH, OW, OY, OR, AR, AA
/NG/	stri**ng**, a**ng**er

*These allophones can be doubled.

TABLE 7 ALLOPHONE ADDRESS

Octal Address	Allophone	Sample Word	Duration	Octal Address	Allophone	Sample Word	Duration
000	PA1	PAUSE	10ms	040	/AW/	Out	250ms
001	PA2	PAUSE	30ms	041	/DD2/	Do	80ms
002	PA3	PAUSE	50ms	042	/GG3	Wig	120ms
003	PA4	PAUSE	100ms	043	/VV/	Vest	130ms
004	PA5	PAUSE	200ms	044	/GG1/	Guest	80ms
005	/OY/	Boy	290ms	045	/SH/	Ship	120ms
006	/AY/	Sky	170ms	046	/ZH/	Azure	130ms
007	/EH/	End	50ms	047	/RR2/	Brain	80ms
010	/KK3/	Comb	80ms	050	/FF/	Food	110ms
011	/PP/	Pow	150ms	051	/KK2/	Sky	140ms
012	/JH/	Dodge	100ms	052	/KK1/	Can't	120ms
013	/NN1/	Thin	170ms	053	/ZZ/	Zoo	150ms
014	/IH/	Sit	50ms	054	/NG	Anchor	200ms
015	/TT2/	To	100ms	055	/LL/	Lake	80ms
016	/RR1/	Rural	130ms	056	/WW/	Wool	140ms
017	/AX/	Succeed	50ms	057	/XR/	Repair	250ms
020	/MM/	Milk	180ms	060	/WH/	Whig	150ms
021	/TT1/	Part	80ms	061	/YY1/	Yes	90ms
022	/DH1/	They	140ms	062	/CH/	Church	150ms
023	/IY/	See	170ms	063	/ER1/	Fir	110ms
024	/EY/	Beige	200ms	064	/ER2/	Fir	210ms
025	/DD1/	Could	50ms	065	/OW/	Beau	170ms
026	/UW1/	To	60ms	066	/DH2/	They	180ms
027	/AO/	Aught	70ms	067	/SS/	Vest	60ms
030	/AA/	Hot	60ms	070	/NN2/	No	140ms
031	/YY2/	Yes	130ms	071	/HH2/	Hoe	130ms
032	/AE/	Hat	80ms	072	/OR/	Store	240ms
033	/HH1/	He	90ms	073	/AR/	Alarm	200ms
034	/BB1/	Business	40ms	074	/YR/	Clear	250ms
035	/TH/	Thin	130ms	075	/GG2/	Got	80ms
036	/UH/	Book	70ms	076	/EL/	Saddle	140ms
037	/UW2/	Food	170ms	077	/BB2/	Business	60ms



PIN FUNCTIONS

Pin	Signal	Functions
A 1	V_p	4.6 to 7.0V DC regulated
B 2	S7	Address Bit 7 (MSB)
C 3	RESET	Initializes Module (Active Low)
D 4	S_0	Address Bit 0 (LSB)
E 5	S_1	Address Bit 1
F 6	S_2	Address Bit 2
H 7	S_3	Address Bit 3
J 8	S_4	Address Bit 4
K 9	S_5	Address Bit 5
L 10	S_6	Address Bit 6
M 11	$\overline{\text{BUSY}}$ (SBY)	Indicates Module is Speaking (Active Low)
N 12	$\overline{\text{STROBE}}$ (ALD)	Loads the 8 address bits
P 13	GND	Power Supply Ground
R 14	$\overline{\text{LOAD}}$ (LRQ)	Load next address
S 15	SPK	Audio Output

ELECTRICAL CHARACTERISTICS

Maximum Ratings*

Temperature Under Bias .... 100° C
Storage Temperature .... −55° to +100° C
Voltage on any pin with respect to GND .... −0.3V to +12.0V

*Exceeding these ratings could cause permanent damage to the device. This is a stress rating only and functional operation of this device at these conditions is not implied. Operating ranges are specified in Standard Conditions. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

Data labeled "typical" is presented for design guidance only and is not guaranteed.

Standard Conditions (unless otherwise stated):
Operating Temperature $T_A = 0°$ C to +70° C
DC CHARACTERISTICS

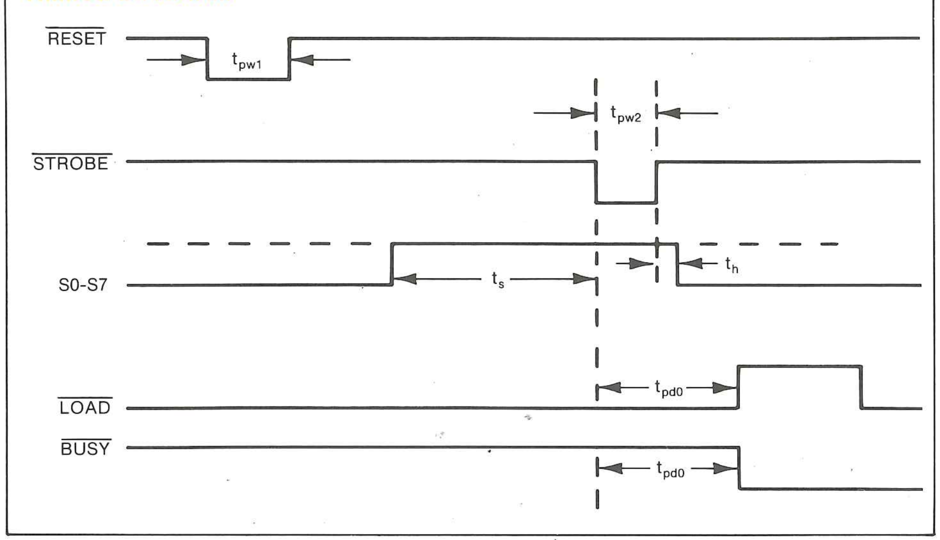
Characteristic	Sym	Min	Typ	Max	Units	Conditions
Supply Voltage	V_P	4.6	—	7.0	V	
Supply Current	I_P		100		mA	Audio Off
S0-S7, Strobe Low Voltage	V_{IL2}	0.0	—	0.6	V	
S0-S7, Strobe High Voltage	V_{IH1}	2.4	—	VP	V	
Reset Low Voltage	V_{IL2}	0.0	—	0.6	V	
Reset High Voltage	V_{IH2}	3.6	—	VP	V	
Busy, Load Output Low Voltage	V_{OL}	0.0	—	0.6	V	0.72mA
Busy, Load Output High Voltage	V_{OH}	3.5	—	VP	V	−50μA
Audio Output	A_P	—	—	200	mW	8Ω Load

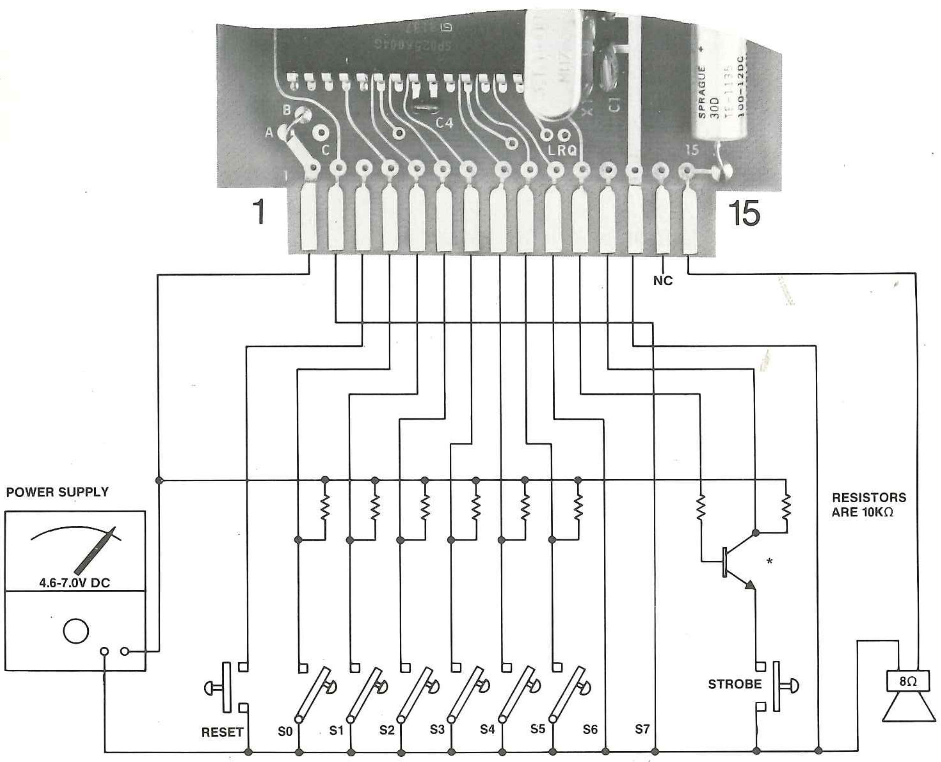
Positive current indicates current into module.
Negative current indicates current out of module.

Standard Conditions (unless otherwise stated):
Operating Temperature $T_A = 0°$ C to +70° C
AC CHARACTERISTICS

Characteristic	Sym	Min	Typ	Max	Units	Conditions
$\overline{\text{RESET}}$	t_{pw1}	100	—	—	μs	
$\overline{\text{STROBE}}$	t_{pw2}	200	—	—	ns	
S0-S7 Set Up	t_s	450	—	—	ns	
S0-S7 Hold	t_h	0	—	—	ns	
Busy 1	t_{pd0}	—	—	300	ns	
Busy 2	t_{pd0}	—	—	300	ns	

TIMING DIAGRAM





*Transistor is used to debounce the STROBE pushbutton switch. If a boundless switch is used, STROBE input pin 12 may be tied directly to switch.

The circuit above demonstrates the simplicity of the interface and illustrates the phonetic sound generation. It would not be practical to attempt word generation with this circuit.

PROCESSOR CONTROLLED INTERFACE

The VSM2128 is easily interfaced to a parallel port on any computer or microprocessor system. The TRS-80™ Model 1 personal computer is used as an example of this operation.

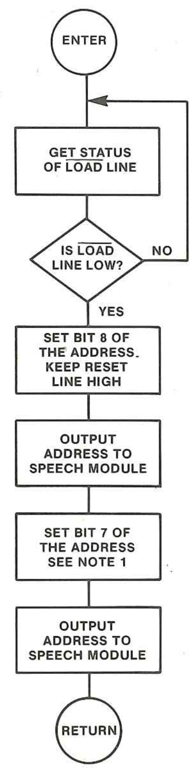
VSM2128 TO TRS80 INTERFACE

VSM2128 CONN.		TRS-80™ LINE PRINTER CONN.		ADDITIONAL CONNECTIONS
1	V_p	—		+5V ±5%
2	S7	—		GND
3	RESET	17	DATA 8	—
4	S0	3	DATA 1	—
5	S1	5	DATA 2	—
6	S2	7	DATA 3	—
7	S3	9	DATA 4	—
8	S4	11	DATA 5	—
9	S5	13	DATA 6	—
10	S6	—		GND
11	BUSY	—		—
12	STROBE	15	DATA 7	—
13	GND	2	GND	GND
14	LOAD	21	BUSY	—
15	SPR	—		8Ω SPEAKER

FLOW CHART

This routine interfaces the TRS-80™ Model 1 to the VSM2128 speech module through the parallel printer port.

The routine waits for the LOAD line to go low (input buffer empty) and then outputs the address of a phrase to the speech module.



Calling program must reset speech module by outputting a low to high transition on pin D_8 of parallel port.

NOTE 1: A high to low transition on D_7 strobes the address into the speech module.

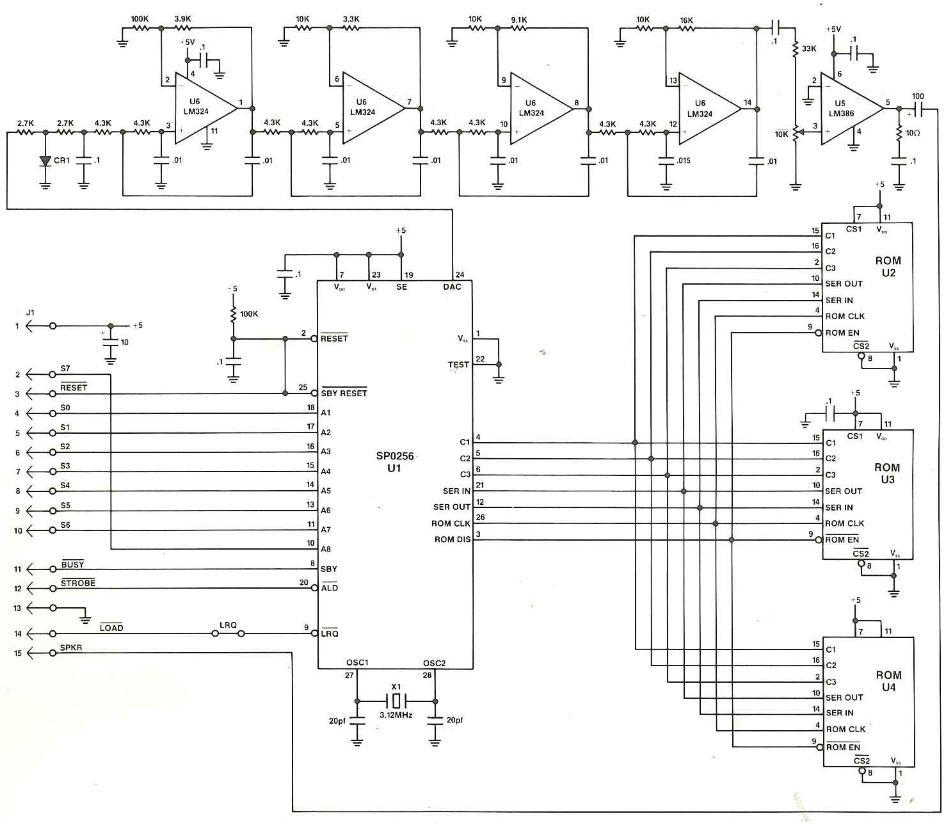
These two statements will reset the speech module.

```
510     POKE 14312,64
520     POKE 14312,192

795     REM OUTPUT ADDRESSES TO SPEECH MODULE
800     XB = PEEK (14312)
810     IF XB < 128 THEN GO TO 800
820     XA(XI) = XA(XI)+128
830     POKE 14312, XA(IX)
840     XA(XI) = XA(XI) +64
850     POKE 14312, XA(XI)
860     RETURN
```

XA: Array — stores the addresses
XI: Index to XA array
XB: Busy yes/no

VSM2128 SPEECH MODULE



NOTES:

1. J1 mates to Amphenol connector 225-21521-401 (117)
2. All capacitor values are μf except where noted.
3. Board supports General Instrument speech ROMs SPR16, SPR32, and SPR128 in any combination.

NOTES