

SC-01A Speech Synthesizer and Related ICs

Last updated 12/14/13, dg



The Votrax SC-01A Speech Synthesizer is a phoneme synthesizer of the early 1980's. It is capable of unlimited English speech using a stream of phoneme codes as input. All transitions between phonemes are handled automatically. (It thus stands in contrast to most other synthesizers of that time--see [below](#).) It was based on work by Richard Gagnon, and later was enhanced to become the SC-02, more widely known as the [SSI-263](#). This page has some resources that might be useful to the person studying or using the SC-01A.

Wikipedia has additional historical information at en.wikipedia.org/wiki/Votrax

(See the bottom of the page for a brief set of [terms](#) that might be useful.)

The speech it produces is of quite high quality for the time (and for a single-chip IC). It rates reasonably well in intelligibility, but is not very natural sounding. 64 different phonemes are available, and 4 levels of intonation.

The SC-01A (nor SC-01) is not available on a regular basis from any known suppliers—what's left is NOS (New, old stock). The chip was last made in the late 1980's, and SSI told us then that due to the types of processes used in this analog chip, it was unlikely to be made again--too complicated to transfer the technology. (However, we never had really good communication with SSI, so take this with a grain of salt.)

Brief comments on the SC-02 / SSI-263 / "Artic 263"

The SSI-263 (or whatever you prefer calling it) is quite an upgrade from the SC-01; it has many more control registers, a different package and pinout, and so on. As far as I know, it has the same analog formant synthesis core as the SC-01, and is also capable of intelligible (but not necessarily natural sounding) speech. "Last buy" was likewise many years ago, so any supplies are NOS.

See below for SSI-263A datasheet and user guide.

The bottom line? Single-chip *analog* formant synthesis is likely down for the count. There are some recent offerings (DSP based), though, that could fill the niche of single-chip phoneme synthesizer--see below under Modern Alternatives.

Data Sheets

Note about audio output from a reader named Cobey: "I wanted to let you know that there are errors on the "9 page recommended" datasheet. If the chip is connected as shown in Fig. 11 or Fig. 13, the audio output will not operate correctly and either be heavily distorted or silent. The "10 page poor" datasheet seems to be a revision which has some extra information including how to calculate clock frequency, proper termination of TP3 and corrected Fig. 11 and 13 (ignoring the scan error in fig. 11)."

I scanned in a copy of the SC-01 data sheet. It is in PDF format, but in fact it is entirely bit-mapped graphics (i.e., it has not been converted to actual text). The resolution is 300 DPI. It is readable for the most part, though some small heavy fonts are hard to read. By the way, after running into a dead-end trying to get permission from Votrax to post this data sheet (they sent me to Artic Technologies, also in the Detroit area, who never got back to me), I noticed the explicit permission granted on the first page to reproduce this data sheet. So feel free to redistribute as you like.

Update--12/07: Bob Grieb was kind enough to furnish an excellent scanned copy of the datasheet. It is scanned at 600 dpi, I think, and the original is loads better than anything else I've seen. I've kept the old copies here as well.

Raymond Weisling kindly forwarded a scan of a better copy of an SC-01 datasheet. This is saved in JPEG format, and I've

also packaged it as a PDF, but again the PDF is just a wrapper around graphic images. I've also included links to the jpg files themselves in zip files if you'd prefer to have them that way. (By the way, check out Raymond's Four Letter Word and other Nixie-tube creations at: <http://www.zetalink.biz/produ.html>)

Files:

Recommended: SC-01 datasheet, 9 pages, good original, PDF: [sc01.pdf](#) (588kB)
 SC-01 datasheet, 9 pages, good original, 150 dpi/JPEG. PDF: [sc01jpeg.pdf](#) (829kB) JPEG/ZIP: [sc01jpeg.zip](#) (805kB)
 SC-01 datasheet, 10 pages (last is application circuits), poor original, 300 dpi/TIFF. PDF: [Poor10PageSc01.pdf](#) (627kB)

If you have an SC-01A datasheet, or related material that could be posted, please let me know.

Reactive Micro (<http://www.reactivemicro.com/>) hosts a copy of the SSI-263 data sheet at:
<http://www.downloads.reactivemicro.com/Public/Apple%20II%20Items/Hardware/SC-02-aka-SSI-263/Datasheet/>

Phonetic Speech Dictionary for the SC-01 Speech Synthesizer: Dave of www.riana.com has posted a PDF of the little dictionary showing SC-01 phoneme sequences for some common words:

<http://www.riana.com/electronics/sc01/index.html>

SC-01 vs. SC-01A

What is the difference? Klatt's review paper below states there is a quality difference between the two. Can anyone go into more detail on the differences?

Answer from Jonathan Gevaryahu: "The SC-01 and SC-01-A have been decapped and the internal roms read out (mid-2007). The difference between the SC-01 and SC-01-A: the 01-A had some of the parameters changed for a few phonemes, ostensibly to increase sound quality and to remove some DC bias from the output."

What happened to SSI?

I'm a little unsure, but it sounds like Texas Instruments acquired a lot of SSI's assets in 1996, according to this press release: <http://www.ti.com/corp/docs/press/company/1998/98005.shtml>

So some SSI elements became part of TI's Storage Products Group (<http://www.ti.com/sc/docs/products/storage/index.htm>).

However, I was in touch (April 2006) with that TI group, and they did not have any information on the 263 (or other chips). They thought that Teridian Semiconductor Corporation (<http://www.tsc.tdk.com/>) might have retained the speech products. However, my email to them remains unanswered. Anyone have additional information?

Resources

For the theory of operation, see the following patents:

Voice Synthesizer, Richard Gagnon, 3,908,085, 9/23/75

Speech Synthesizer Responsive to a Digital Command Input, Richard Gagnon, 3,836,717, 9/17/74

Voice Synthesizer, Mark Dorais (assigned Federal Screw Works (=Votrax)), 4,128,737, 12/5/78

Voice Synthesizer, Carl Ostrowski (assigned Federal Screw Works), 4,130,730, 12/19/78

Integrated circuit phoneme-bases speech synthesizer, Carl Ostrowski and Bertram White (assigned Federal Screw Works), 4,433,210, 2/21/84

See also Gagnon, R. T. (1978). "Votrax Real Time Hardware for Phoneme Synthesis of Speech," Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-78, 175-178. (Many thanks to Eric Smith for finding this citation!)

From "Talking Terminals," David M. Stoffel, Byte, September, 1982: "The Votrax VSA and VSB synthesizers seem quite similar with respect to their phoneme production, but the FSST-3, which uses the VSA, definitely sounds inferior; whether this is an artifact of the VSA synthesizer or poor audio amplification, I don't know. You may wonder why none of these products uses the new Votrax SC-1A (sic) integrated circuit, which is less expensive. The single quantity cost of the VSB is about \$800, while the SC-01A is \$70. But there are two major reasons why the SC-01A is not used. The speech-rate and pitch controls are both dependent on the same clock signal or timing circuit, affecting the ease with which intelligible

speech may be produced. Also some people are concerned about the acceptability of the SC-1A's (sic) sound quality. Only scientific performance measures can determine which Votrax synthesizer is ultimately more intelligible. (For a description of an application using the Votrax SC-01A speech-synthesizer chip see Steve Ciarcia's article on page 64 in this issue.)" See <http://www.lindenreport.com/stoffel/talk.html>

For the best overview of speech synthesis up to the late 80's, see Klatt, Dennis, "Review of Text-to-Speech Conversion for English," *Journal of the Acoustical Society of America*, 82:3, September 1987, p 737-793. This is an *excellent* overview, and includes (pg 756) a brief description of the SC-01 that starts as follows:

"Apparently oblivious to all of the prior research detailed earlier, a man experimenting in his basement workshop, Richard Gagnon, designed a synthesis-by-rule program that eventually resulted in the Votrax SC-01 chip. ... It is a remarkable device for the price."

No mention of Mozer (of National Semiconductor's speech synthesizer, if I'm remembering correctly), but this is the best article to start with if you want an introduction to the synthesis of speech.

Important note: Klatt's paper is now online! See this and other papers at http://www.mindspring.com/~ssshp/ssshp_cd/ss_home.htm

Prochnow, Dave, *Chip Talk: Projects in Speech Synthesis*, Tab Books, Blue Ridge Summit, PA: 1987. ISBN is 0-8306-1912-7 (hard cover) and 0-8306-2812-6 (paperback). Hardcover is Tab # 2812. A hobbyist-oriented book on speech synthesizers circa 1987, including

- Votrax SC-01A (analog formant)
- General Instruments SP0256-AL2 (CTS256A-AL2)
- National Semiconductor DT1050 Digitaltalker (Mozer)
- Silicon Systems SSI 263 (analog formant)
- Texas Instruments TMS5110A (LPC)
- Oki Semiconductor MSM5218RS (ADPCM)

as well as stand-alone systems. Schematics, pin-outs, construction hints. Not much detail on theory. All but certainly out of print, but that's what inter-library loans are for!

For a deep account of one text-to-speech system (the basis for some of the best speech synthesizers until perhaps recently), see the book *From Text to Speech: The MITalk System*, by Allen, Hunnicutt, and Klatt, Cambridge University Press, 1987. (The pseudo-code in the back is not, however, without a number of errors and omissions.) The parameter-to-speech part of MITalk is detailed for the most part in Klatt's "Software for a Cascade/Parallel Formant Synthesizer," *J. Acoust. Soc. Am.*, 67:3, March 1980, pg 971-995.

See http://www.mindspring.com/~ssshp/ssshp_cd/ss_home.htm, the Smithsonian Speech Synthesis History Project, which includes audio of a variety of synthesizers, the Klatt paper, and personal recollections. See especially the chronology of Votrax's speech products.

Obviously, newsgroups such as comp.dsp, comp.speech.*, comp.arch.embedded, and comp.robotics.*--and their associated FAQ's--are valuable resources.

See also <http://www.robotprojects.com/voice/voice.htm>, by Scott Savage, which has some interesting links on speech synthesizers.

(7/2/02) Tom McClintock notes the following: "One item of interest regarding the SC-01a. The 'PinMAME' developers have incorporated SC-01a emulation into their pinball simulations. The source code includes digital representations of all the phonemes. Pretty cool stuff, but complete and accurate emulation is not quite there. Check out the source: http://pinname.retrogames.com/release/pinname_112_1_src.zip"

Bob Paddock wrote up a nice list of links at <http://www.chipcenter.com/circuitcellar/june00/c0600rp42.htm>
Kevin Horton has reverse engineered a number of Votrax-based speech synthesizers (VSL, Type and Talk, PSS, etc.): <http://www.kevtris.org>

Alex Bettarini has a website where you can have an actual SC-01a produce speech from what you type in: <http://real-votrax.no-ip.org/> "Users can program their own phonemes or choose words from a dictionary (in continuous expansion). It's possible to choose different clock frequencies to get a wide range of voice pitches. Finally, each user can download the generated audio file. Of course it's all free. What I am working on now is implementing a text-to-speech algorithm built into the system, as well as providing musical functionality using MIDI karaoke files."

Applications

Ciarcia, Steve, "Build a Low-Cost Speech Synthesizer Interface," Byte, June 1981, p 46.

Ciarcia, Steve, "Build an Unlimited-Vocabulary Speech Synthesizer," Byte, September 1981, p 38.

Ciarcia, Steve, "Build the Microvox Text-to-Speech Synthesizer, Part 1," Byte, September 1982, p 64.

Ciarcia, Steve, "Build the Microvox Text-to-Speech Synthesizer, Part 2," Byte, October 1982, p 40.

(See <http://members.tripod.com/werdav/t2smicrv.html> for the article above.)

Ciarcia, Steve, "Talk to Me: Add a Voice to Your Computer for \$35," Byte, June 1978, p 35.

Ciarcia, Steve, "Build a Third-Generation Phonetic Speech Synthesizer," Byte, March, 1984, p 28. (SSI-263)

Note that some of these articles are collected in "Ciarcia's Circuit Cellar" volumes I, II, and III. Volume I covers 9/77-11/78, II covers 12/78-6/80, and III covers 7/80-12/81.

Vernon, Peter, "Add Speech to Any Computer with the Compuvoice Computer Speech Synthesizer,"; Electronics Australia, October 1982, pg 72-78. Complete description of a SC-01-based circuit board (including PCB artwork), Centronics parallel interface. (Thanks to Mark Best.)

Moffat, Tom, "The Chatterbox -- Computer Voice Synthesizer", Electronics Today International (Australia), January 1985, pg 74-81. (Thanks to Mark Best down under.)

Technical details of Gottlieb pinball machines, some of which use an SC-01A chip:

http://www.ionpool.net/arcade/gottlieb/technical/sound_boards.html

The Type N Talk manual: <http://members.tripod.com/werdav/txtospm1.html>

Intex Talker (uses SC-01A): <http://web.inter.nl.net/hcc/davies/txtospin.html>

Where it was used

One approach to finding SC-01A's is to buy used devices that have this chip in them (e.g., eBay). The following include an SC-01A:

- Intex Talker (<http://web.inter.nl.net/hcc/davies/txtospin.html>)
- Votrax Type N Talk
- Microvox Text-to-Speech Synthesizer (<http://members.tripod.com/werdav/t2smicrv.html>)
- Heath Hero robots (but the robot is worth way more than the SC-01!)
- Voice Box, Atari voice synthesizer by The Alien Group
- VS100 synthesizer board for TRS-80 Model III (<http://ripsaw.cac.psu.edu/~mloewen/Oldtech/Tandy/>)
- SPD 125, a speech synthesis board for the Apple II, by Speech Design (ca. 1982) (thanks to P. Jansen ([photo](#)))
- RB5X Robot (thanks to Ethan Dicks)--still being sold (\$3,500), apparently, as of June 2007: <http://www.rbrobotics.com/Products/RB5X.htm>. While they do list an SC-01A in their parts list, the availability is shown as "call/email".
- Braid Speech Synthesizer (ca. 1985), used a 6502 CPU with 2 x 4K EPROMs (2732). (Thanks to Richard Hutchinson)
- Project "Orac" (after a talking computer in a British TV show "Blakes Seven") used an SC-01 as a annunciator for alarms in an electricity distribution control center (in Australia?). (Thanks to Mark Best)

- Gottlieb arcade game hardware (Gottlieb speech/sound assembly A3), used in games such as Qbert, Qbert Qubes, Krull, and Curveball. (Thanks to Elmar Trojer) For more information, see qbert.me

(what others am I missing?)

Brief comments on other single-chip synthesizers of the past:

The late 70's to mid 80's saw a number of speech synthesis chips developed. Below are brief comments to place some other chips of that era in relation to the SC-01.

Texas Instruments TMS5100, TMS5220, etc. TI developed a series of synthesizers using LPC (linear predictive coding). LPC is a compact method of encoding speech, and so typically systems using these chips have a limited, pre-specified vocabulary, though in a few instances software and lots of data was used to create unlimited speech systems (e.g., Street Electronics). (LPC is still used as the core of most digital speech compression algorithms, including for digital cell phones.)

General Instruments also produced some LPC-based synthesis chips. (And one chip that supported LPC analysis for speech recognition--the SP1000? Ciarcia had an article on the chip.)

General Instruments SP0256-AL2. According to Sclater (Neil Sclater, *Introduction to Electronic Speech Synthesis*, 1983, Howard W. Sams), a digital ~~formant~~ synthesizer using a serial data stream from special ROMs. Can be paired with the **CTS256A-AL2**, which is a hard-coded PIC7041 microcontroller (this, according to Prochnow) that has a built-in text-to-speech algorithm and ROM with allophones. (Is there any transitioning of parameters by the built-in program? Or are allophones just concatenated, which would drastically reduce the quality of the speech? I don't know.) **Update!** See <http://www.primenet.com/~im14u2c/intv/tech/ivoice.html> for a very detailed description of the SP0256 along with C source to simulate it, by Joe Zbiciak. The SP-0256 is in fact a 12-pole LPC synthesizer and not a formant synthesizer. It has a simple microsequencer that can execute a handful of instructions. (Thanks to Eric Smith for the link!)

Philips PCF8200. A digital formant synthesizer, it requires a constant flow of parameters to synthesize speech. Typically these parameters were derived from actual speech, but in theory, you could create these parameters using software (such as Klatt's algorithms) to provide unlimited speech. (Formant parameters are "easily"--and directly--synthesizable from abstract rules; LPC parameters are not easily directly synthesizable from sets of rules.) Essentially, a digital version of the formant filters in an SC-01, but without the transitioning logic found in the SC-01 (such transitions in the SC-01 were generated using analog circuitry).

National Semiconductor DigiTalker (MM54104). A direct waveform encoding/decoding chip set. Uses ROMs with a limited number of words.

Modern alternatives

1. Voice recording chips

Currently available is a line of voice recording chips from **ISD (now Winbond)**. These are even available at Radio Shack and Digikey. However, they are waveform recording devices, so not capable of unlimited speech.

2. Single-chip phoneme synthesizer

(2/22/04) A new pre-programmed PIC that does single chip speech synthesis and sound effects--the **SpeakJet**.

Apparently released within the past two weeks, it accepts serial data in (phonemes) and output a PWM signal that with minimal (2-pole) filtering can be fed to an amplifier and then speaker. Internal oscillator. Seems to run about \$25.

Developed apparently by Magnevation (www.magnevation.com) and Scott Savage (oopic.com) over the past 5 years. I have only heard a few demos. Widely available through robotic supply sources. The interface and command set look very well thought out. This might turn out to be a very nice chip for applications that would normally want to use the SC-01A. According to an email I got from Scott, the SpeakJet does do transitioning between phonemes. If anyone has additional details how this works, what PIC it is (someone guessed an PIC18F1320), etc., let me know. (This conjecture makes sense--the 18F1320 has an 8x8 multiplier, 8k bytes of program space, PWM, and runs about 10 MIPs. This is more than enough to do a stripped-down digital formant synthesizer. A full bore, unoptimized KLATTalk-ish formant synthesizer core will run on a 10 MIPs 16-bit wide chip with MAC.)

Update (12/4/07): Robert Doerr (<http://www.robotworkshop.com/>) just wrote an article in the December issue of Servo Magazine (<http://www.servomagazine.com/>) about using a small microcontroller to translate SC-01 phonemes into

SpeakJet allophones, plus handle the interface signals so you can plug the circuit into a regular SC-01 22-pin socket. If you need to replace an SC-01, but keep the rest of your circuitry intact (e.g., Hero robot), this could be an interesting solution.

Also, Chip Gracey, Parallax founder and the designer of the Propeller chip, has apparently been working on speech synthesis that would run on the Propeller. (See Make magazine volume 10: <http://www.make-digital.com/make/vol10/?pg=78&search=parallax+propeller+speech&u1=texterity&cookies=1>.) Anyone with additional information? If it ran on just one of the eight 32-bit processors (which should be quite realistic), this would be interesting for new embedded applications.

3. Single-chip text-to-speech

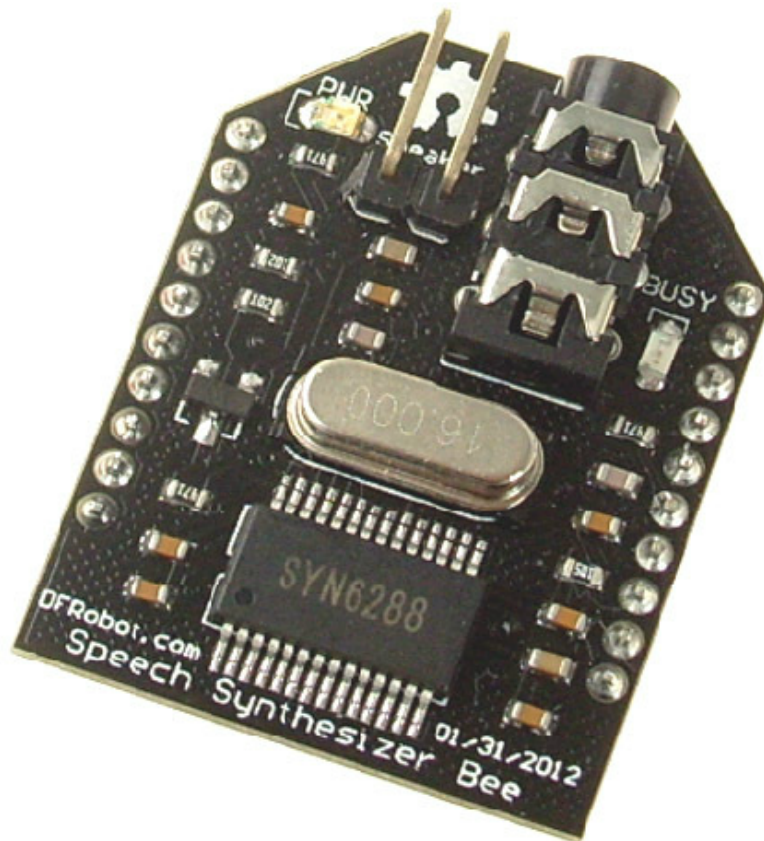
(7/2/02) Robert Doerr points out a newish chip from Winbond, the WTS701, which includes text-to-speech algorithms. http://www.winbond.com/E-WINBONDHTML/partner/b_2_a_5.htm.

(5/14/03) Tom Arnold points out that the datasheet is finally available for the WTS701, along with a live demo (you type in text, get back audio output) From the description, it sounds like it stores speech (using the ISD technology) on chip, concatenating to form the output. (See their FAQ on the page above.) Surface mount package. SPI interface.

Update (7/11/12): Eric Ostendorff pointed out the RoboVoice SP0-512 Text to Speech IC. <http://www.speechchips.com/shop/item.aspx?itemid=22>. About \$25 (\$15.99 on special as of 7/11/12). A pre-programmed dsPIC33FJ64GP802 PIC microcontroller (Microchip), it claims an 800-rule TTS algorithm, built-in DAC, serial communication. Eric has written an [article about the RoboVoice](#) in the November 2012 Servo magazine. Here's a link to a [video he made](#).

Update (7/11/12): Eric Ostendorff also sent a link to a module, the Emic 2 Text-to-Speech Module, \$59.95 from Parallax. It is based on a single-chip TTS from Epson, the V30120, and also has a 32-bit Freescale ColdFire micro—I suppose to make the interface easier? The V30120 embeds the Fonix DECTalk TTS engine, communication is SPI. The package is TQFP 13-64, so not the most DIY-friendly. I see Mouser can order the V30120 in minimum qty 401 in 14 weeks lead time for \$6.17 a piece. This is very interesting! They also appear to make some chips that provide additional features such as MP3 decoding, etc., in addition to the TTS. Languages supported by this chip are US English, Castilian Spanish, and Latin American Spanish.

(12/14/13) Eric Ostendorff sends in another text-to-speech module based on the "SYN 6288" chip. On eBay, about \$30. Here's an example of a module in the XBee form factor:



"Working Voltage :3.3-5V; Interface Type : TTL serial port, default baud rate 9600; Providing speakers Interface; Providing 3.5 headphone jack; Compatible with Xbee socket"

4. Do it yourself DSP

You could also port open-source speech synthesizers to a microcontroller/DSP platform. This is getting easier all the time—you'll need on the order of 5-10 MIPS of 16-bit wide processing (with MAC), and the digital-to-analog output along with at least 32-64 kB of program space plus some RAM. (That sentence used to start "This is non-trivial...", but we have some amazing options these days for microcontrollers—plenty of RAM, 16-32 bit architectures, and high enough clock rates and/or DSP-friendly instructions to do this.)

As for speech synthesis by concatenation ("why not just record all 64 sounds from the SC-01 and string them together as you want?"), see the comment at the bottom of this page.

Sources for various chips

Please note that I do not have any SC-01A's or 263's for sale.

(9/2011) Fred Teer has a few tubes of SC-01A's, NOS, that he is selling. Sounds like around \$25 each, but contact him for a quote: fredteer@yahoo.com

(8/2010) Kevin Keinert offers SC-01A's for \$41 (qty 1) at <http://mysite.verizon.net/res8aiig/ICparts/ICparts.htm>, along with SP1000, SP0250, TMS5200, and others.

(8/2010) Reactive Micro has some SC-01A's and SSI-263P's for \$50 each: http://www.reactivemicro.com/index.php?cPath=1_42

(8/2010) A source for some speech chips (not including the SC-01A): <http://www.speechchips.com/shop/> They sell the SpeakJet, SP0256-AL2, and others.

A few useful terms:

Formant: A peak in the spectrum of speech corresponding to a resonance in the vocal tract. A formant synthesizer uses bandpass filters, typically, to create these resonances. Depending on the type of speech to be synthesized (e.g., female, male), four or usually five or more formants are necessary for reasonable speech.

Phoneme: An abstract sound unit, for example a sound like "eh". The problem is that the actual sound associated with a phoneme depends on its context--the "eh" sound is influenced by what sounds precede and follow it, for example. The SSI-263 has some 60 phonemes it uses to synthesize English speech.

Allophone: The actual realization of a phoneme (i.e., in a particular context). More than one allophone can be associated with a given phoneme.

One more comment...

One idea that seems to come up when discussing the SC-01A (specifically, the lack of availability), is that of creating a software emulation by recording the 60+ phonemes and just concatenating them. It turns out that much of the intelligibility of speech is wrapped up in the transitions between the ideal phoneme sounds ("targets"), so the concatenated speech will sound nowhere near as good as the SC-01, which has internal circuitry to generate transitions. In fact, there's a speech synthesis method based on just concatenating the transitions between phonemes, called diphone synthesis.

Feel free to e-mail me if you have any interesting resources or note errors on this page: [dgrover at redcedar.com](mailto:dgrover@redcedar.com)