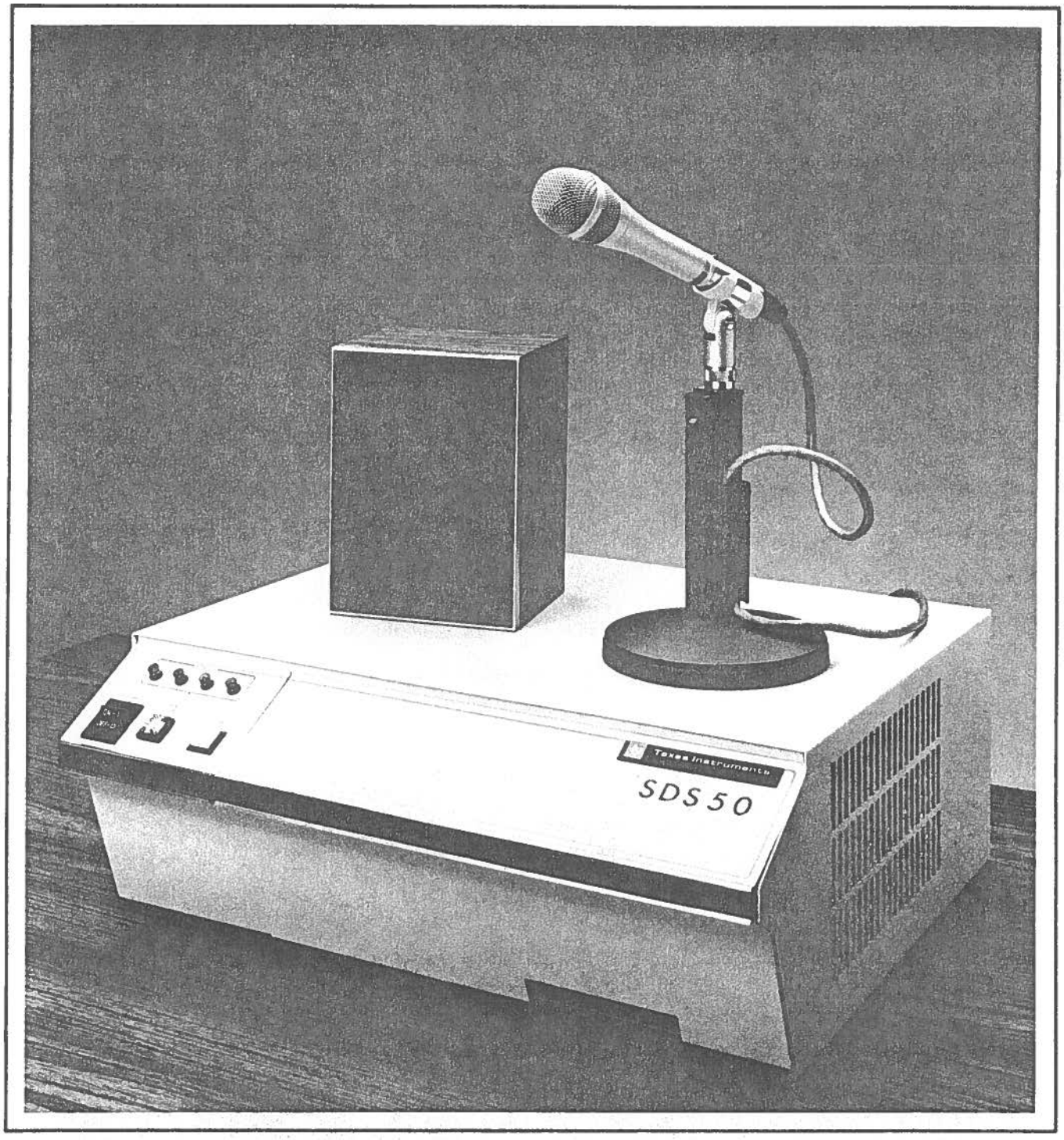
TEXAS INSTRUMENTS

Speech Development System
SDS50



Product Brief

THE PORTABLE SPEECH LAB FROM TEXAS INSTRUMENTS

Machines communicating vocally with humans represents a rapidly expanding market. Applications span virtually every market segment from consumer, to automotive, to telecom, to computer, to industrial uses.

While the reproduction of speech from encoded data has been reduced to a readily available chip set, up until very recently encoding speech and storing it in memory has required a great deal more effort and hardware. Texas Instruments Speech Development System is a self-contained speech analysis system which allows one to listen to data encoded for Texas Instruments voice synthesis processors the instant the person has spoken a sentence.

The necessary hardware has been reduced in size from a mainframe computer to a standard Texas Instruments microcomputer and a number of dedicated signal processing boards.

THE TRADITIONAL METHOD

Before the advent of the Texas Instruments Speech Development System (SDS50), the potential user made speech recordings at the vendor's site. Then an LPC algorithm was executed to encode raw speech which was subsequently edited for use in a product.

This technique required a great deal of mass storage for the digitized, uncoded speech. In addition, it was not until after a considerable period of time after the recording that the encoded speech could be auditioned. If a potential user was not happy with the result, the speaker would have to be recalled, the speech re-recorded and re-processed — a time consuming iteration.

THE NEW METHOD

The Speech Development System dramatically reduces the time, effort, and assets involved in putting speech into silicon, by directly converting the spoken text into "Linear Prediction" coefficients and synthesizing the results immediately to the user. Not only does this ensure that users can iteratively optimize the results obtained, but it also permits the particular words to be encoded at short notice.

SYSTEM CONFIGURATION

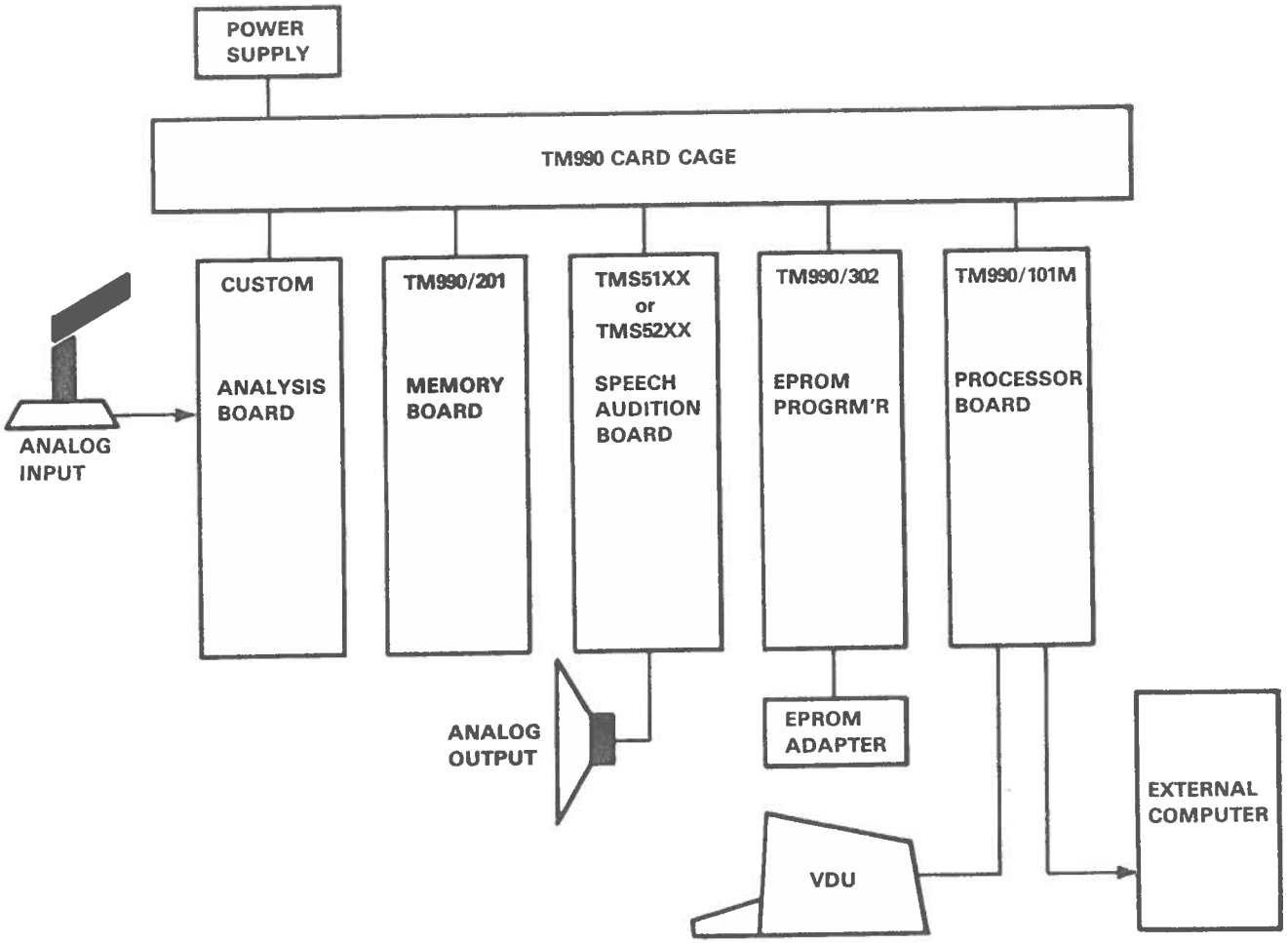
The Speech Development System may be used in three modes:

(1) As a stand-alone demonstration system with only a microphone and a loudspeaker. This is useful for evaluating and training speakers.

(2) As an interactive system with a VDU. This permits speech data to be displayed in terms of the Linear Prediction Codes which are used to regenerate utterances and these can then be edited to improve quality. Further, new sounds can be created in this way. Once the appropriate speech has been captured. The EPROM programmer may be used so that the speech stored in the SDS50 may be downloaded directly into nonvolatile memory for use in an actual product.

(3) As a speech processing front end to a computer. In this way, speech can be transferred to an archive so that when various words or phrases are needed for subsequent refinement by editing or reuse they can quickly be retrieved from an archived file.

The SDS50 not only fills a gap in the speech synthesis development cycle but also forms the basis for speech development laboratories. The unit's ability to do this is due both to its unique hardware design and the implementation in software of a proprietary algorithm.



TECHNICAL DETAILS

a) The Portable Speech Lab contains:
TMS990 Card Chassis
Speech Analysis Board
Synthesis Board for TMS51XX or 52XX VSPs
TM990/302 EPROM Programmer Board
TM990/101 CPU Board
TM990/201 Memory Expansion Board
Loudspeaker
Microphone and preamplifier
Power Supply

b) Data Output:
Program EPROMS, TMS2516, 2532, 2708, 2716 in bit and byte wide formats.

c) Speech Input:
Microphone/600Ω line (tape or cassette)

d) Supports: TMS5110A, 5220, 5220A VSPs.

e) Digital Output:
EPROM byte serial data
bit serial data

f) Analysis:
Either 8 or 10kHz sampling rate
Data rate produced = 1600/2000 bps approx

g) Storage:
Approx 10 seconds of speech

Speech recorded on the SDS50 can be programmed into an EPROM. This is done by using the EPROM programmer and a standard terminal attached to the system via an RS-232C interface. The system provides a menu of commands among which is the "(P)ROGRAM EPROM" command.

The (A)NALYSE ANOTHER PHRASE command allows the terminal operator to set up the SDS50 to record a new phrase. Once the phrase is recorded, it can be heard by striking the R key, which is the (R)EPLAY PHRASE command. In a Lab, the next likely step would be to use the (L)IST or (E)DIT PHRASE PARAMETERS command.

When either command is invoked, the LPC parameters are displayed on the screen as in diagram 1.

The energy parameter (ENGY) is listed first, followed by the gain parameter (CE), the pitch parameter (P) and the reflection coefficients (K1 through 0).

For a novice, the parameters that are simplest to edit are energy and pitch. For instance, if speech has been captured and, later, there is some syllable which should be emphasized more strongly, a user can increase the energy parameter to do this. At any time during the editing process the phrase can be replayed to see what effect the editing has had. In a similar way, the pitch of the recorded speech can be altered to give a phrase a more pleasing or consistent tone.

The reflection coefficients can also be edited and the SDS50, with its instant playback capability, is thus a useful tool for making engineers familiar with the intricacies of LPC. Once a user is familiar with the effect of these coefficients, any sound which the human vocal tract is theoretically capable of producing can be created.

Diagram 1

£	ENGY	CE	P	K1	K2	K3	K4	K5	K6	K7	8	9	0
001	0152	02	00	00	09	10	07						
002	0102	02	00	00	09	09	07						
003	0057	01	00	07	09	09	07						
004	0029	05	15	19	03	05	07	07	07	08	4	3	3
005	0037	05	15	17	03	07	08	08	07	06	4	4	4
006	0034	05	16	18	05	05	06	07	08	08	3	3	3
007	0034	05	17	17	05	06	06	07	08	07	3	4	3
008	0026	05	17	20	04	06	05	06	06	07	4	3	4
009	0147	10	16	21	18	10	03	08	07	05	4	3	3
010	0490	12	16	19	06	07	07	11	10	09	6	2	1
011	0567	12	16	17	11	08	09	10	08	10	5	0	2
012	0788	13	16	16	13	10	09	09	08	11	5	1	2
013	0469	12	16	17	11	11	08	09	06	10	4	1	1
014	0422	12	17	17	10	12	07	09	05	11	4	1	1
015	0388	12	17	17	10	11	08	08	05	11	3	1	1
016	0332	11	17	16	12	10	08	07	07	11	3	1	1
017	0328	11	17	17	10	09	09	07	07	09	5	1	1
018	0332	11	18	17	11	10	07	08	07	12	5	2	1
019	0401	11	18	17	11	09	09	08	07	11	6	2	1
020	0408	11	18	16	11	10	09	08	07	10	6	2	1
021	0347	11	18	17	08	11	09	09	08	08	6	3	1
022	0283	10	18	17	07	11	09	09	07	08	6	4	0
023	0195	09	18	16	11	11	07	06	08	10	6	3	1
024	0123	08	18	16	09	11	07	08	05	09	5	2	1
025	0111	08	18	16	10	11	06	07	04	11	5	4	2
026	0103	07	18	16	13	10	05	04	04	10	6	4	2
027	0071	07	18	16	08	09	08	08	05	07	5	4	4
028	0072	07	17	17	07	11	07	07	05	07	4	4	4
029	0075	08	17	18	08	11	07	05	03	07	5	3	4
030	0086	07	17	17	12	12	06	04	03	08	5	3	4
031	0057	07	17	17	08	11	07	05	04	06	4	4	4
032	0078	02	00	08	15	10	05						
033	0037	01	00	16	13	10	04						
034	0039	01	00	14	10	10	07						
035	0039	01	00	14	09	09	06						
036	0036	01	00	15	12	09	05						