[redacted]

JET PROPULSION LABORATORY

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA

JPL 5030-76

VOICE DATA ENTRY SYSTEM DEMONSTRATION

April 30, 1977

R. L. Sohn

Prepared for

National Criminal Justice Information and Statistics Service

LAW ENFORCEMENT ASSISTANCE ADMINISTRATION

U.S. DEPARTMENT OF JUSTICE

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91103

ACKNOWLEDGEMENTS

Several persons participated in the feasibility tests summarized in this report. Acknowledgements are made to Carl Geuy and Dave Malan, Consultant, who supported the data file query demonstration at the Fresno Police Department; to Ed Carri for the JPL voice recognizer radio access tests; and to Gabe Moschetti and Shirar Bhanji of Scope Electronics for their general support. We also wish to thank Chief Hal Britton, Deputy Chief Jim Packard, Sgt. Earl Kaundart and others of the Fresno Police Department for their support and cooperation.

CONTENTS

1. INTRODUCTION AND SUMMARY

This report describes feasibility experiments with Voice Data Entry Terminal Systems (VDETS) for law enforcement applications. Two applications were investigated:

- Use of VDETS as a keyboard replacement for mobile and base station data entry terminals. This enables the operator or field officer to query a computer based data without the assistance of a relay operator, and without the necessity of typing instructions via a keyboard.
- Access to a VDETS unit over a mobile/portable radio channel. This would permit a field officer to query a computer-based data file from a mobile/portable radio with VDETS equipments located at the base station.

The resuls of these tests are summarized in this report.

Background

These investigations were prompted by two observations. First, mobile digital terminals have proved to be quite successful for making data file queries; query rates have increased markedly when field units equipped with these terminals have been given direct access to data files without the need to relay requests through base station operators. However, while field personnel generally approved of the automatic data base query capability, certain disadvantages were frequently pointed out: the mobile digital terminals were too bulky, and aggravated an already overcrowded condition in the front seat of patrol cars; keyboards were difficult to access; and displays were sometimes hard to read. Voice data entry devices could eliminate the need to type instructions, an advantage while driving or during car stop operations, and would eliminate the space taken up by the keyboard.

Secondly, VDETS technology has advanced very rapidly in the past few years, and the recent dramatic reductions in the sizes and costs of digital equipments have greatly improved the outlook for applying this technology to law enforcement communications.

Voice Recognizer Systems

A number of "second generation" voice data entry systems are currently on the market, based on 1970 voice analysis techniques and component hardware.

Developments are now underway to microminiaturize these systems but none are available in today's market. Three companies, Scope Electronics, Threshold Technology and Perception Technology, currently market "second generation" voice recognition systems, all supported by minicomputers. The systems are designed for discrete word recognition, as opposed to continuous speech recognition, have basic vocabularies of about 40 to 50 words, require 16K of core, and cost from $10,000 to $20,000. Dialog, Inc., offers a somewhat more sophisticated voice recognizer in the $35,000 - $45,000 price range, based on a PDP11 minicomputer. The Dialog system was designed primarily for continuous speech applications, and is somewhat "over-designed" for mobile law enforcement communications.

Heuristics, Inc., of Los Altos, recently announced an interesting micro voice recognizer designed for the hobby market, selling for approximately $1,000. The voice feature extractor consists of a single board containing a signal generator, 3 bandpass filters, and an A/D converter; the board is priced at about $250. The feature extractor is interfaced to an Intel 8080 microprocessor to identify utterances. An identifier system based on the Heuristics hardware is currently in use at the University of San Francisco for hearing impairment research. The system has a vocabulary of 16 words. While we did not have an opportunity to test the suitability of this system for a law enforcement mobile digital communications application, the device warrants further consideration for a number of reasons. It is small, inexpensive and requires minimal power (190 milliwatts).

Data File Query

A Scope Electronics, Inc., system was leased to perform the data base query and direct RF access feasibility tests. One version of the Scope System is being modified for an Army forward observer application in which the observer inputs data to the fire control center via the voice data entry terminal that is accessed over an RF link from a distance of up to several miles. Hence, some background data were available for the direct RF access tests.

A data base query capability was demonstrated by interfacing a VDETS unit with a Kustom MCT-10 mobile keyboard/display terminal at the Fresno Police Department. The MCT-10 was linked to the Department's computer system via a dedicated digital radio channel, although both the MCT-10 and VDETS units were physically located at the base station. The computer routes data file queries received from MCT-10 or other terminals through a switcher to local data files, and via the California Law Enforcement Telecommunications System (CLETS) to files in Sacramento, and to the NCIC files in Washington. The MCT-10 has proved to be highly effective for data file queries, has a full keyboard, uses digital communications links, and hence is well suited to comparing keyboard versus voice data entry.

The VDETS was supplied with a vocabulary of 48 words, including the 36 alphanumerics, 4 function words (REGISTRATION, LICENSE, etc.,), and 8 control words (TRANSMIT, etc.,); this set covers all MCT-10 operations except some cursor manipulations.

The demonstration was conducted with good success, and queries were made on stolen vehicles, licenses, and persons; the software was modified to accomodate administrative messages, and this function was demonstrated as well. With sufficient training, VDETS accuracy was well within acceptable limits. Responses to data file queries were obtained in 5 to 15 seconds. The system was operated by Department personnel, and several commented that they would like to have such a capability when operating field units.

Direct RF Access Tests

Tests conducted by Scope Electronics at Washington, D.C. and by JPL at Pasadena to access a VDETS unit over a radio channel using mobile/portable radios were less successful. For the Washington tests, a VDETS unit located at the base station was interfaced by cable to the District's radio network, such that incoming voice messages could be routed directly to the VDETS system. Attempts were then made to access VDETS from portable/mobile radios at some distance from the base station. These attempts were unsuccessful because of the excessively noisy signals input to the VDETS unit. Because of limited time and funding, no attempts were made to identify and remove the noise sources.

More successful attempts at direct RF access were made at JPL, using a VDETS unit and Motorola PT200 portable radios. Prior to field access tests, the distortion and response characteristics of a pair of PT200's were measured at several signal input levels. The PT200's were then interfaced by cable to the VDETS unit and access was attempted with the transmitting unit in a screen room. These tests were successful over a range of signal input levels. It was noticed, however, that the Nova minicomputer that supports the VDETS unit radiated energy to the radio receiver, causing significant interference to the voice signals. This interference does not detract from VDETS operation under high signal-to-noise conditions, as with the conventional microphone; when signal-to-noise ratios are low, as with portable radios operating under field conditions, the computer noise cannot be tolerated. By separating the receiver from the computer, moderately successful access was achieved from distances up to 1½ miles. It was concluded that radio access could be achieved if precautions are taken to screen the computer from the RF units, otherwise the low signal-to-noise ratio environment associated with mobile/portable radios is too seriously degraded.

Conclusions

It was concluded that voice recognition technology is currently available for application to law enforcement mobile communications. A good arrangement for this purpose is one in which the voice feature extractor is located in the patrol car and the supporting computer in the base station. In this way, consistently accurate input to the voice recognizer is assured, with minimal loading on the RF channel. A display screen should be used in the patrol unit for convenience. A small, low-cost, low-power voice recognizer such as Speechlab has not been tested, but is illustrative of the equipment needed.

It is recommended that LEAA continue monitoring and testing the feasibility of voice recognizer devices. Minimal funding is required for this purpose.

2. VOICE RECOGNITION TECHNOLOGY

Research in automatic voice recognition has been in progress for many years. It received a notable impetus in the 1970's from programs sponsored by the Advanced Research Projects Agency (ARPA), particularly in the study of automatic analysis of syntax and semantics. Accumulated knowledge from analysis/synthesis studies provied the basis for initial design of automatic speech recognizers; this technology has been channeled into several applications, and systems with vocabularies of several hundred words (with designated speakers) are tractable.

Generally, four processing elements are contained in a limited vocabulary voice recognition system: transducer/microphone, preprocessor, feature extractor, and a classifier. Early attempts at automatic recognition either deleted the feature extractor function, or utilized a simplified form of template matching. However, experience with crude template matching soon brought out its limitations, primarily the inability to distinguish between small variations in individual speech samples of a particular word.

The preprocessor samples the voice utterance and measures those parameters that are used by the feature extractor and calssifier in identifying the utterance.* It is desirable to obtain the spectral energy distribution and its variation with time, with sufficient resolution in both the time and frequency domains to record the information-bearing properties in both domains. This information is then analyzed either by means of direct analog spectrum analysis, by fast Fourier transform, or by linear predictive coding techniques. The specific technique used and complexity involved depend on the accuracy required, the size of vocabularly specified, and whether continuous speech or discrete words are being processed. In any case the spectrum analysis is only the first step in the processing of the speech signal; considerable additional processing is required to detect and extract the significant features of the utterance.

The feature extraction function varies considerably from one implementation to the next. Although considerable mathematical formulation has been developed for speech recognition, no general theory exists that can preselect the information-bearing portions of the speech signal. In the Threshold Technology system, which could be classified as "intermediate" in complexity, the spectral shape and its time derivatives are derived in the feature extraction process.

---

*T.B. Martin, "Practical Applications of Voice Input to Machines." Pages 487-501. IEEE Proceedings. V.64, No. 4, April 1976. (Special issue on man-machine communication by voice.)

Combinations and sequences of the measurements are processed to produce a set of 32 acoustic features, including the word boundary, several broad-class features such as vowels, consonants, pauses, bursts, etc., and a number of phonetic event features. These processing functions are accomplished principally by hard wired units, which makes it possible to achieve real-time processing.

The classifier function is performed in software using a micro or minicomputer. In the Threshold Technology system, the 32 encoded features and their times of occurence are stored in a short-term memory. The utterance is divided into 16 equal time segments, and the features reconstructed on a normalized time base. The pattern-matching logic compares these features to stored reference patterns and determines the best fit for a word decision. A total of 512 bytes of information (32 features and 16 time segments) are required per utterance.

Since the voice input system is adaptive, it must be trained for individual speakers and words. Usually several training passes are required to develop an adequate set of reference patterns; these can be stored on tape and entered into the processor when a given speaker wishes to use the system.

The Scope Electronics system uses a somewhat simpler feature coding technique. The spectrum analyzer divides the input audio spectrum into 16 frequency bands that cover the useful frequency range.* The resulting 16 analog signals define a power spectrum that constitutes the feature for word classification. The 16 continuous signals are sampled at 100 Hz and converted to digital form with 8 - bit resolution. A coding compressor reduces the spectral data to a 240 - bit pattern per word, giving a compression ratio of about 100. The comparation and training functions are similar to those described above.

---

* J. J. Kalinowski; J. C. Brown, S. G. Bhanji, M. G. Hooten, and J. W. Preusee, "Application of Discrete Word Recognition and Response to Multiuser Tactical Communications."

The "Speechlab" system produced by Heuristics, Inc., of Los Altos, is a simplified version of the Scope Electronics device.* Three frequencies only are sampled, plus the average frequency of the utterance. Samples are obtained every 10 millisecs, and the duration of the utterance measured. The duration is divided to give 16 evenly spaced parameters from the three frequency bands and zero-crossing information. This yields 64 bytes of information (16 parameters from each of 4 bands), which are compared to similar measurements made during training passes. The "template" with the smallest difference from the sample is selected as the answer. The Speechlab vocabulary is limited to about 16 words. It can be used with an 8080 microprocessor, and is available in computer hobby shops for $250 (less computer, I/O and memory boards).

As noted in the Introduction, a voice data entry system can be used to advantage for law enforcement communications as a keyboard replacement for mobile digital terminals, such a the Kustom MCT-10. Since the keyboard itself has a limited number of characters (about 50), the voice data entry systems also need not have an extensive vocabulary. The Threshold Technology and Scope Electronics systems are both adequate in this regard. For this application the classifier unit and supporting computer would be located at the base station, so that a single unit could support a number of mobile terminals; each terminal would require preprocessor and feature extractor/compressor devices. The Speechlab vocabulary is probably too restricted, but an "intermediate" capacity device that has a vocabulary larger than that of Speechlab but smaller than those of the Threshold/Scope units should be adequate. Hopefully, the cost of such an in-car unit would be no more than $1,000; the unit could be packaged in the terminal, and use the present visual display and radio interface devices.

The development of a unit with these performance characteristics should be encouraged.

---

* H. Erea and J. Reykjalin, "Introducing Speechlab- The First Hobbyist Vocal Interface for a Computer." Polular Electronics, May 1977.

3. VDETS FEASIBILITY TESTS FOR MOBILE DIGITAL COMMUNICATIONS

Mobile keyboard/display terminals have been installed in patrol cars by a number of law enforcement agencies to give officers in the field direct access to computer-based data files, such as lists of stolen vehicles, other stolen property, and wants or warrants on persons. To use the mobile terminal, the officer types out brief instructions, such as a vehicle license number, a REGISTRATION function key, and a TRANSMIT key. The message is modulated onto the mobile communications channel (either shared voice or dedicated digital channel), and transmitted to the base station computer/switcher. The query is then sent to local data files, and to state and national files if desired. Responses are sent directly to the mobile terminal, usually within 10 to 20 seconds, and displayed to the officer. The mobile terminal enhances officer safety because he can more readily check to determine if a stopped car has been stolen, or the driver wanted.

Voice data entry devices could eliminate the need to type instructions, an advantage while driving or during car stop operations, and would eliminate the space taken up by the keyboard.

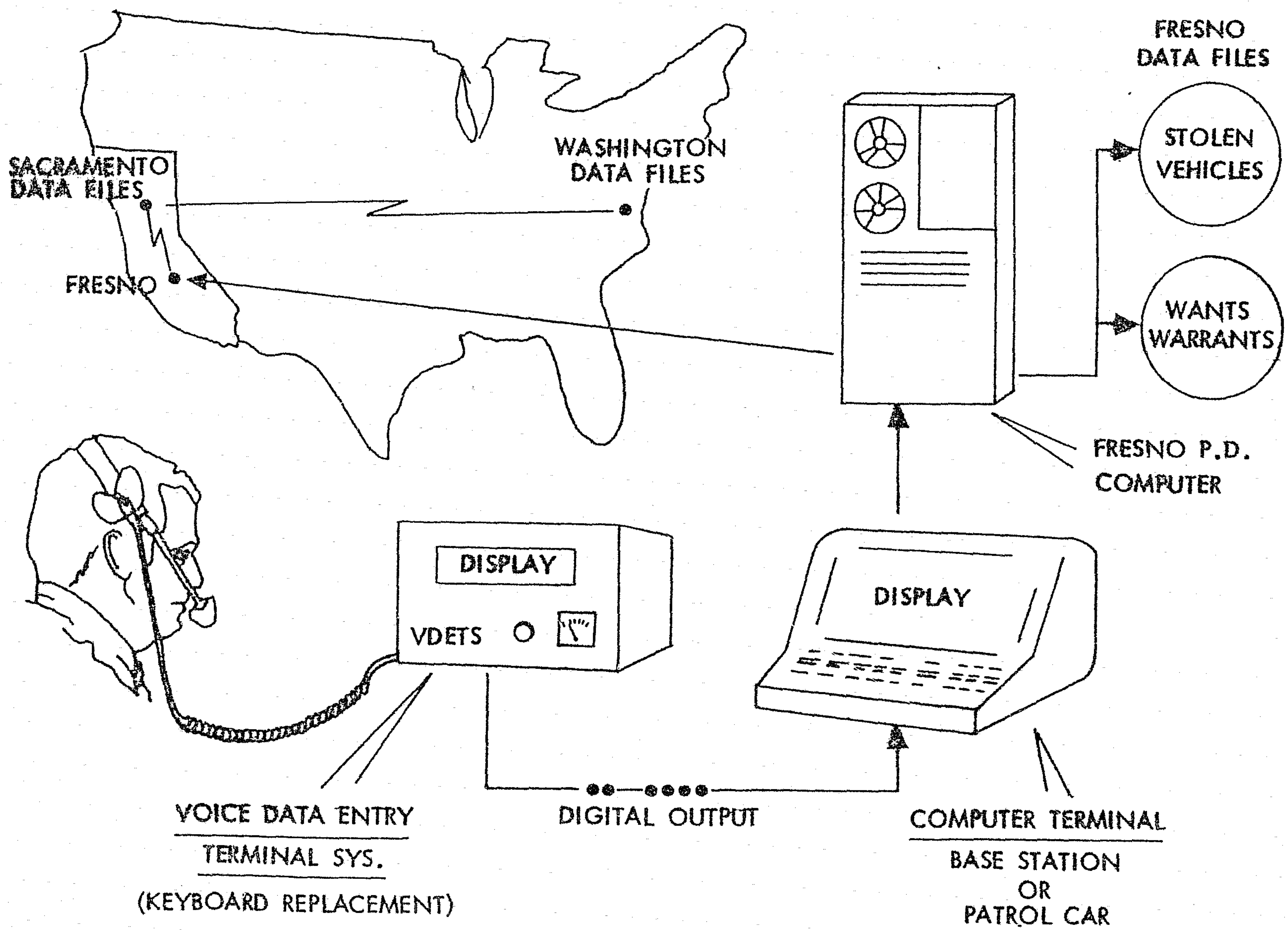
The feasibility of using a VDETS system as a keyboard replacement for a mobile terminal was demonstrated at the Fresno Police Department (Fresno, CA), on April 22, 1977. The Fresno PD was originally contacted for this demonstration because the agency had recently installed a computer-aided dispatch (CAD) system, and has equipped a number of patrol cars with Kustom Mobile Communications Terminals (MCT-10), which are linked to the CAD computer by a dedicated digital communications channel. The base station computer accesses local data files in response to MCT-10 requests, and can switch queries to state files at Sacramento via the California Law Enforcement Telecommunications System (CLETS); queries can also be directed to the NCIC files using the CLETS/NCIC networks (see Figure 3-1).

For the demonstration, the VDETS unit was interfaced with and MCT-10, which accessed the CAD computer over the dedicated digital communications channel; the MCT-10 and VDETS were physically located at the base station, however.

The vocabulary for the demonstration was chosen to exactly duplicate that available via the keyboard, which consists of the alphanumeric characters (26+10), function keys to indicate the type of query (registration, license, person, administrative, etc.), and two control keys for TRANSMIT and DISPLAY.

FIGURE 3-

VOICE DATA TERMINALS FOR LAW ENFORCEMENT



A left cursor shift was also added for ease in correcting input errors. The MCT-10 Operator Manual was followed closely to assure compatability with normal keyboard manipulations and user procedures.

The VDETS unit frequently uses a syntax structure for the purpose of improving accuracy by narrowing the number of utterances over which the classifier searches; for example a given utterance could indicate to the classifier that only digits need be searched for classification, and so on. Since the use of syntax structure would detract from officer convenience, syntax structure was not used. The results justified this approach since a high accuracy was achieved without the syntax structure, due in part to the relatively small vocabulary used (see Table 3-1).

A software program was developed based on the vocabulary and operational procedures presented in the MCT-10 Operators Manual.* The program was successfully developed in a relatively short time, and checked out at the Scope Electronics facility prior to the demonstration.

The VDETS/MCT-10 interface board was designed to convert the VDETS output to emulate that of the MCT-10 keyboard. This involved a serial to parallel conversion and necessary support logic. This unit was successfully installed with minimal debugging.

A schematic of the VDETS demonstration is shown in Figure 3-2. The VDETS NOVA 2/10 computer is programmed to operate in two modes: TRAIN and OPERATIONAL.

In the TRAIN mode, the user is prompted through each word of the vocabulary as displayed on the Self-Scan unit. As the user pronounces a vocabulary word, VDETS digitizes this response and stores it in a preselected location in memory. For the purpose of this demonstration 5 to 10 vocabulary training passes were found to be sufficient for good speech recognition.

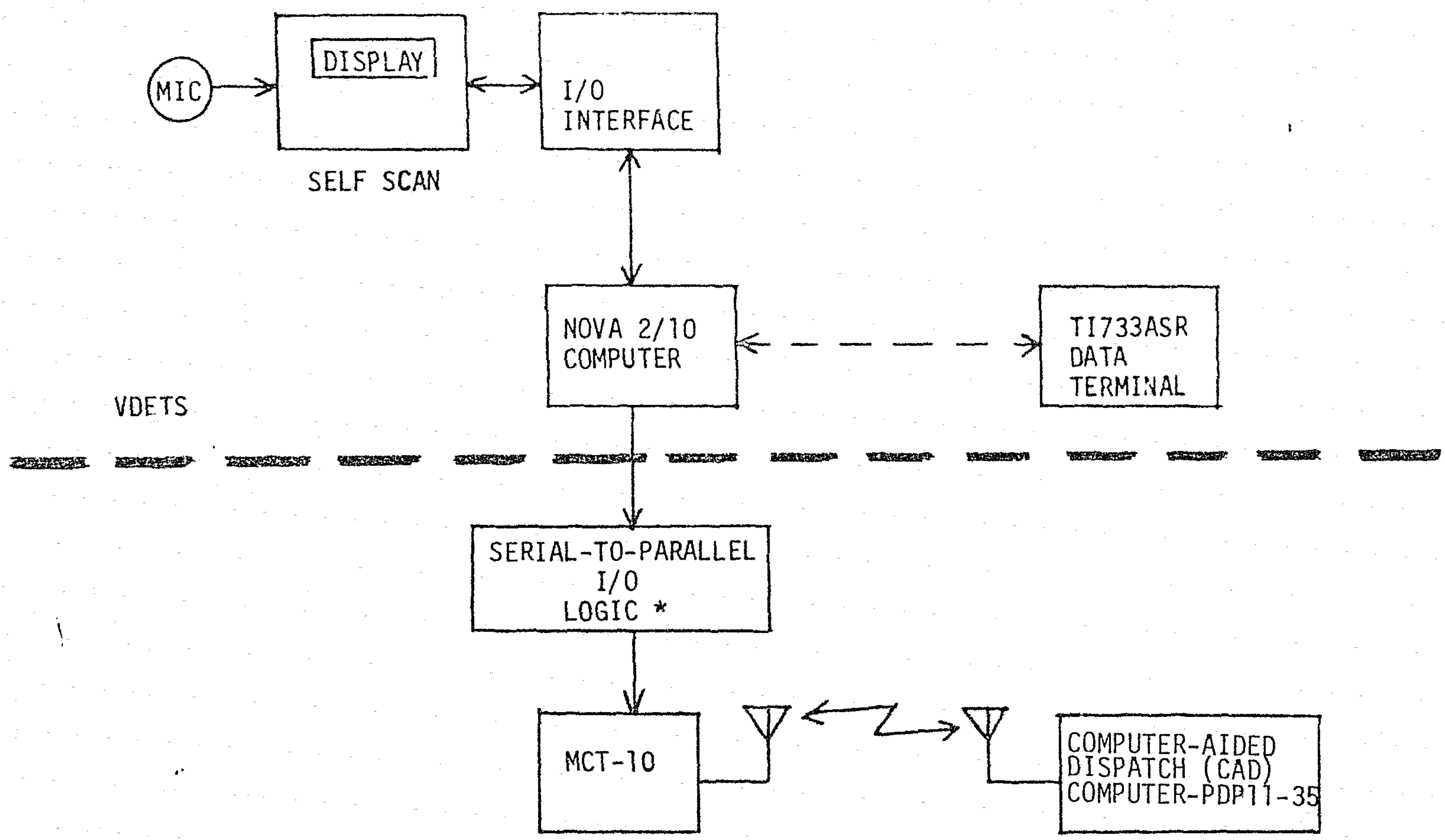
In the OPERATIONAL mode, VDETS digitizes each vocabulary utterance and compares it to all vocabulary words stored during the TRAIN mode. When a match is found, the ASCII character code is sent to Self-Scan and to the MCT-10 and displayed on each. If a match is not found, "PLEASE REPEAT" is displayed on the Self Scan. In an actual deployment of such a system, the Self Scan display would not be necessary once the TRAIN phase has been completed. Each user's voice pattern, as stored in the NOVA 2/10 computer, would be recorded on a cassette at the end of the TRAIN phase. Thereafter, when the officer goes on duty, he would playback the cassette to restore his voice patterns in the computer.

---

* "Fresno Police Department Kustom MCT-10 Operators Manual."

Table 3-1 VDETS VOCABULARY

VDETS TRAIN	OPERATION	ASCII (MCT-10) OCTAL	COMMENTS
CORRECTION	CORRECTION	47	Correct Inquiry Message
REGISTRATION	REGISTRATION	41	Function Code F1
LICENSE	LICENSE	42	Function Code F2
PERSON	PERSON	43	Function Code F3
SPECIAL	SPECIAL	44	Function Code F7
TRANSMIT	TRANSMIT	46	Initiate Data File Inquiry
DISPLAY	DISPLAY -	45	Display Data File Retur
ADAM	A	1	
BOY	B	2	
CHARLES	C	3	
DAVID	D	4	
EDWARD	E	5	
FRANK	F	6	
GEORGE	G	7	
HENRY	H	10	
IDA	I	11	
JOHN	J	12	
KING	K	13	
LINCOLN	L	14	
MARY	M	15	
NORA	N	16	
OCEAN	O	17	
PAUL	P	20	
QUEEN	Q	21	
ROBERT	R	22	
SAM	S	23	
TOM	T	24	
UNION	U	25	
VICTOR	V	26	
WILLIAM	W	27	
XRAY	X	30	
YELLOW	Y	31	
ZEBRA	Z	32	
ZERO	Ø	60	
ONE	1	61	
TWO	2	62	
THREE	3	63	
FOUR	4	64	
FIVE	5	65	
SIX	6	66	
SEVEN	7	67	
EIGHT	8	70	
NINE	9	71	
PERIOD	.	56	
SLASH	/	57	
COMMA	,	54	



*HARDWARE FABRICATED
FOR THIS DEMONSTRATION

FIGURE 3-2. VDETS DEMONSTRATION SCHEMATIC

The TI 733ASR Data Terminal consists of a TTY and cassette record/playback equipment. It is only used during the TRAIN mode, and for recording/playback of the NOVA 2/10 program/voice patterns; it is not used during the OPERATIONAL mode.

Use of the VDETS vocabulary in the transmission of data requests/messages is as follows. The request/message is constructed using vocabulary words "ADAM" through "COMMA", followed by one of the four function codes "REGISTRATION" through "SPECIAL." The user then transmits the request/message by speaking the word "TRANSMIT."

The message response from the CAD computer is indicated at the MCT-10 by an audible "BEEP." At this time, the user may display the message response on the MCT-10 by speaking the word "DISPLAY." The user can clear the MCT-10 display in preparation for another data request/message by again speaking the word "DISPLAY."

At any time prior to speaking the word "TRANSMIT," the user may correct a constructed data request/message by speaking the word "CORRECTION." This moves the MCT-10 cursor back one space in the message, at which time the user may enter the correct vocabulary word.

The demonstration was conducted with good success, and queries were made on stolen vehicles, licenses, and persons; the software was modified to accomodate administrative messages, and this function was demonstrated as well. With sufficient training, VDETS accuracy was well within acceptable limits. Responses to data file queries were obtained in 5 to 15 seconds. The system was operated by Department personnel, and several commented that they would like to have such a capability when operating field units.

It is recommended that feasibility tests be conducted with the simplified feature extractor board (Speechlab) discussed in Chapter 2. An ideal arrangement would have the feature extractor board installed in the patrol unit and interfaced with the MCT-10 terminal (with or without keyboard). Voice messages would be coded by the feature extractor and transmitted through the MCT-10 to the base station for identification processing, and returned to the MCT-10 for display. Thus, a keyboard replacement capability is obtained through use of a relatively inexpensive but accurate feature extractor in the mobile unit and a single CPU at the base station. Channel loading is minimal, hence the basic advantage of digital communications is not sacrificed. Conceivably, this approach could be extended to portable radios.

4. DIRECT ACCESS TO A VOICE RECOGNIZER FROM MOBILE/PORTABLE RADIOS

Tests were conducted by JPL at Pasadena and by Scope Electronics at Washington, D.C., to access a VDETS unit via a radio channel using mobile/portable radios. This capability would permit field officers to query computer-based data files via mobile/portable radios with VDETS equipments located at the base station, thus giving the officer's direct, automated access to remote data files without the need to install voice recognition equipments in all field units. Such an arrangement would not reduce the channel loading experienced by an all-voice radio network, but would constitute the least expensive voice data entry system. The test programs are described in the following sections.

4.1 JPL TESTS

The JPL tests involved direct access of a VDETS unit via an RF link using Motorola PT200 portable radios. Prior to field access tests, the distortion and response characteristics of a pair of PT200's were measured at several signal input levels. The PT200's were then interfaced by cable to the VDETS unit and access was attempted with the transmitting unit in a screen room. These tests were successful over a range of signal input levels. It was noticed, however, that the Nova minicomputer that supports the VDETS unit radiated energy to the radio receiver, causing significant interference to the voice signals. By separating the receiver from the computer, moderately successful access was achieved from distances up to 1½ miles.

The specific objectives of the tests were:

- Characterize a typical land mobile radio channel for its voice transmission capability.
- Demonstrate in a controlled laboratory test the ability to interface and communicate with a VDETS unit via a radio link.
- Demonstrate in a field test the ability to communicate with a VDETS unit via a radio link.

The approach taken to achieve these objectives included the following steps:

- Calibrate a controlled radio path in terms of receiver quieting.
- Measure the distortion and frequency response of a typical Land Mobile Radio (LMR) pair for several signal input levels.
- Interface LMR with a VDETS unit. Using specified words, measure the ability of VDETS to correctly respond as receiver input level is reduced.

- From remote locations, transmit specified words by a mobile transmitter for reception and processing by a VDETS unit. Determine the quality of detection as the transmission range increases.

The transceivers used were Motorola Model PT 200 hand held units operating at 40.820 MHz with a nominal output of 1.4 watts. The Voice Data Entry Terminal System, Model 1000, was made by Scope Electronics. The VDETS unit is supported by a Nova 2/10 minicomputer. A chart of the control program vocabulary syntax structure is shown in Figure 4-1.

The equipment block diagrams are shown in Figure 4-2 for the three test phases. For the calibration and laboratory demonstration tests, the transmitting transceiver was located in a screen room.

<u>Calibration Phase</u>

Figure 4-3 shows the receiver output noise voltage with unmodulated carrier input as a function of attenuator setting. This curve is referred to as a receiver quieting curve, and shows that quality reception is obtained with attenuator settings of less than 140 db. When the signal is attenuated beyond this point, the receiver output noise can mask the intended voice output.

Figure 4-4 shows the measured frequency response of the transceiver pair for three different values of attenuator setting. A relatively poor response was obtained over the voice band of interest.

Spectrum analyzer photographs of the receiver output are shown in Figure 4-5 with a 1 kHz tone modulating the transmitter. Note that harmonics are quite evident, but are, in fact, considerably reduced from the output at 1 kHz and should not be of concern. Backgound noise level increases as the attenuator setting is increased so that the signal is almost obscured at a setting of 150 dB.

<u>Laboratory Demonstration Phase</u>

When using the controlled radio path to communicate with the VDETS, error free recognition was achieved down to attenuator settings of 135 db. Acceptable recognition was also achieved at higher db settings but only if the VDETS unit was retrained to recognize the speaker's voice. At attenuator settings of 165 dB, some words could still be recognized but errors were frequent.

During this test, it was noted that the minicomputer associated with the VDETS was radiating sufficient energy to be received by the transceiver and cause interference to the voice signals. Shielded enclosures are recommended for future operations of this type.

<u>Field Tests</u>

For this test, the transmitter was separated by distances up to 2 miles from the receiving site. The test vocabulary was then transmitted and the

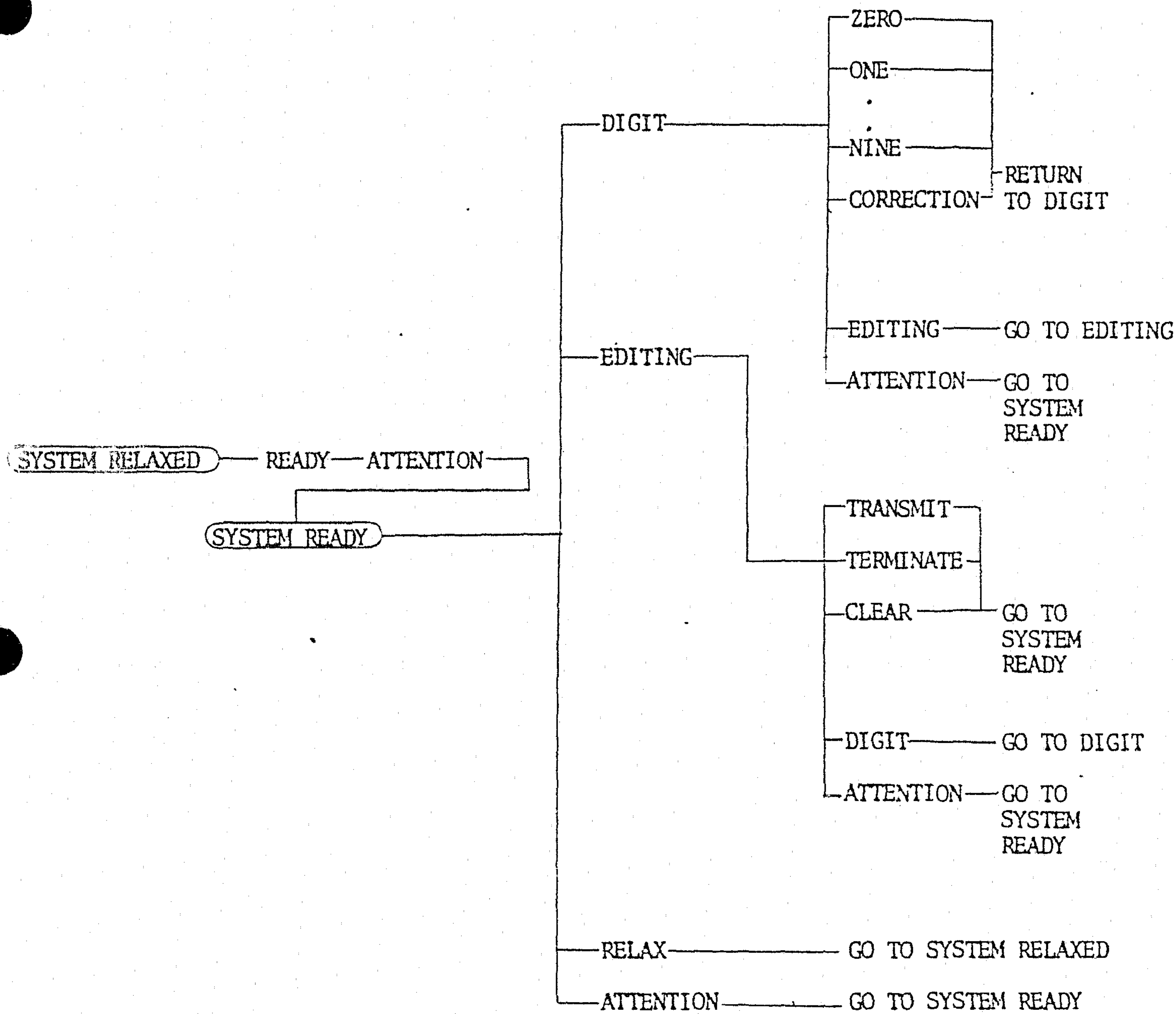
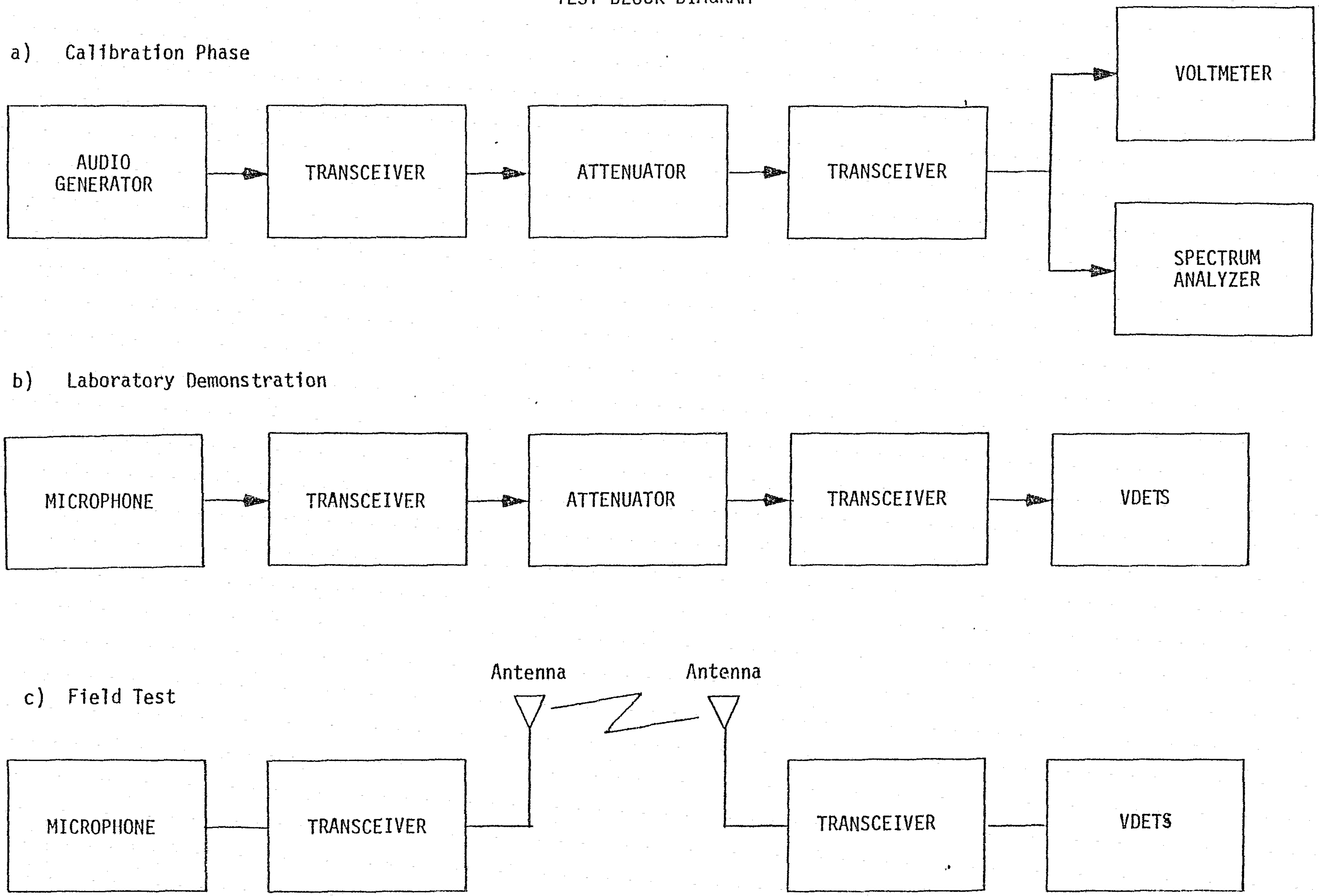


Figure 4-1 VDETS Control Program
Vocabulary Syntax structure

Figu[illegible]-2

TEST BLOCK DIAGRAM



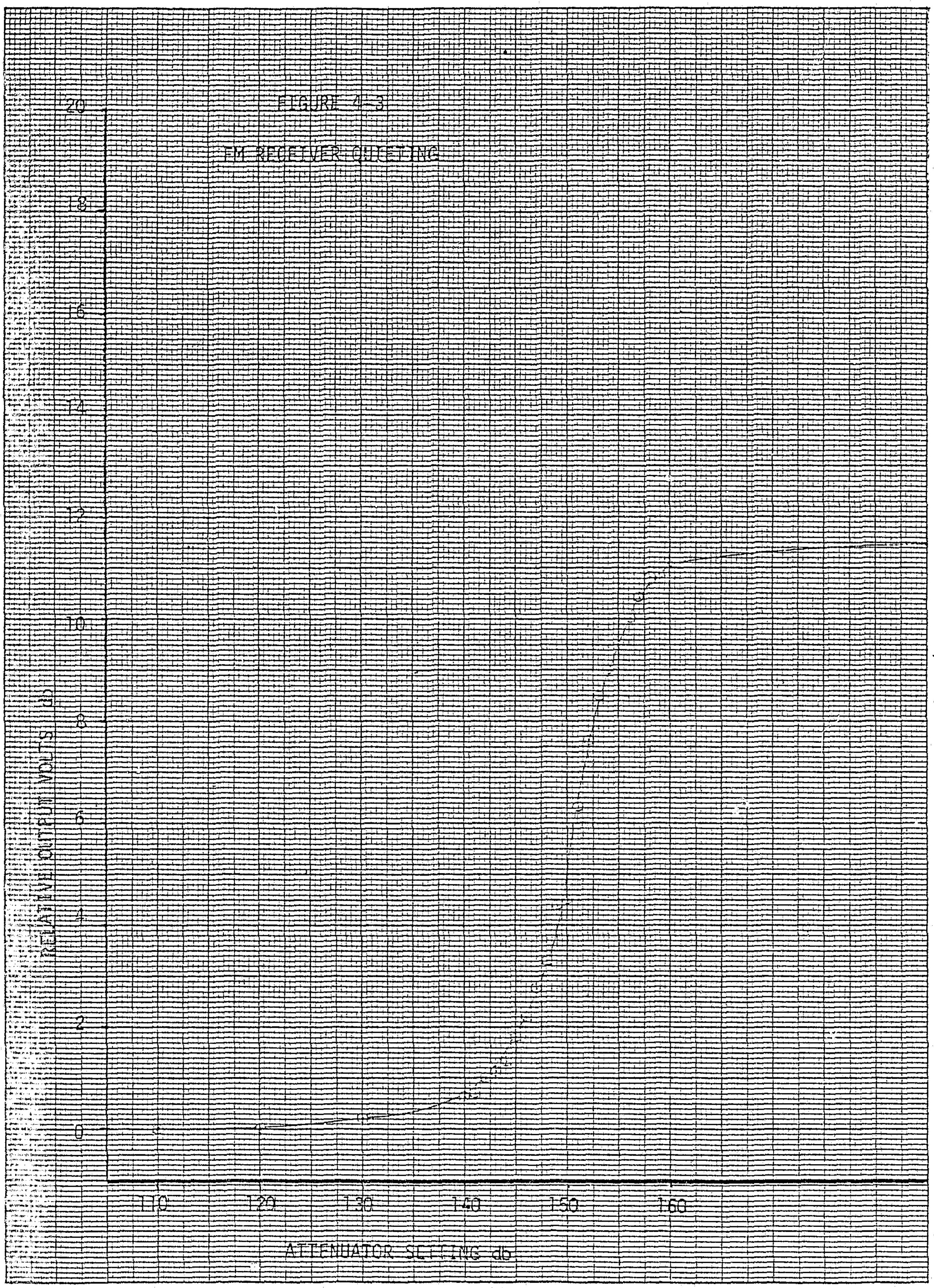


FIGURE 4-3
FM RECEIVER QUIETING
RELATIVE OUTPUT VOLTS db
20
18
16
14
12
10
8
6
4
2
0
110
120
130
140
150
160
ATTENUATOR SETTING db

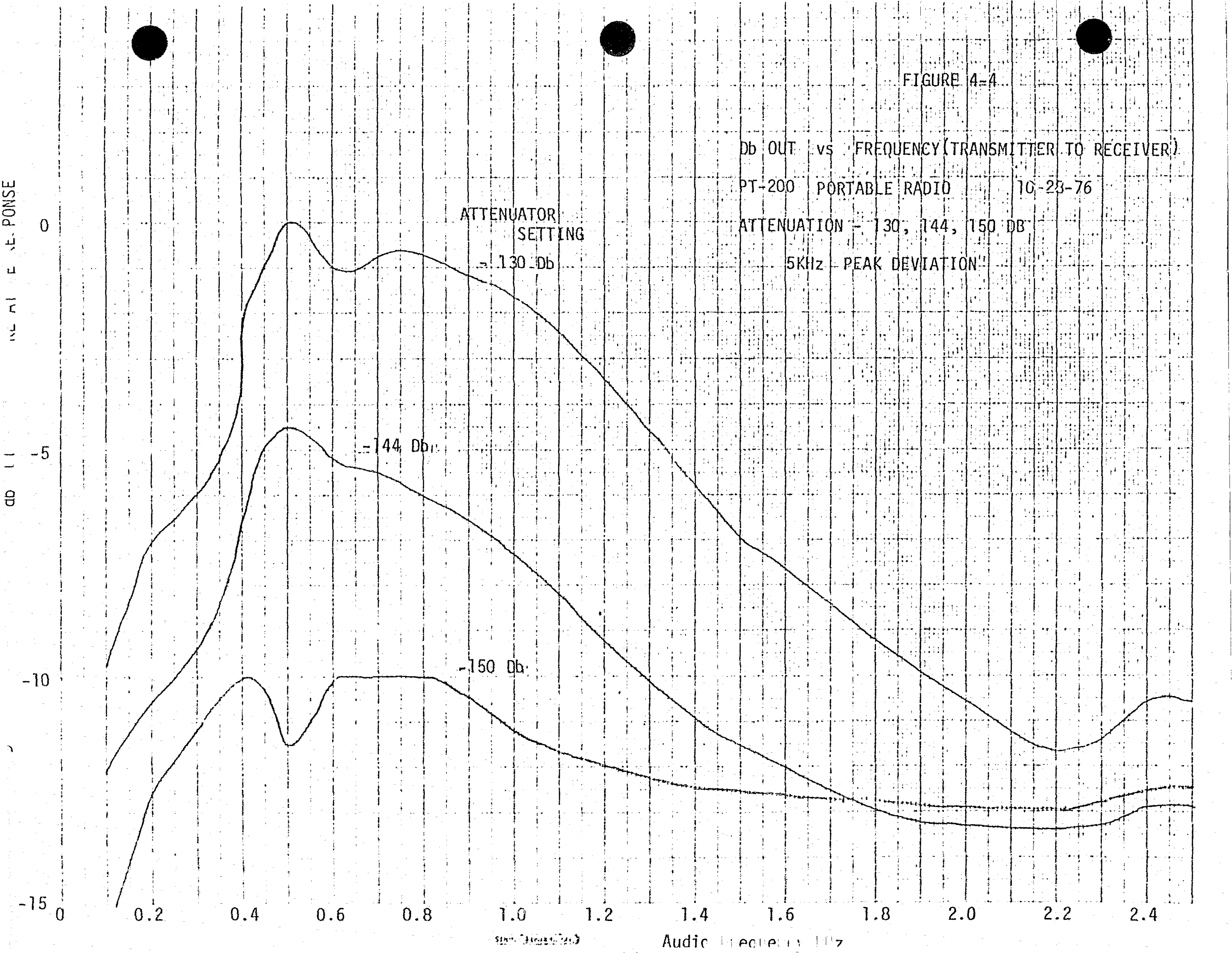
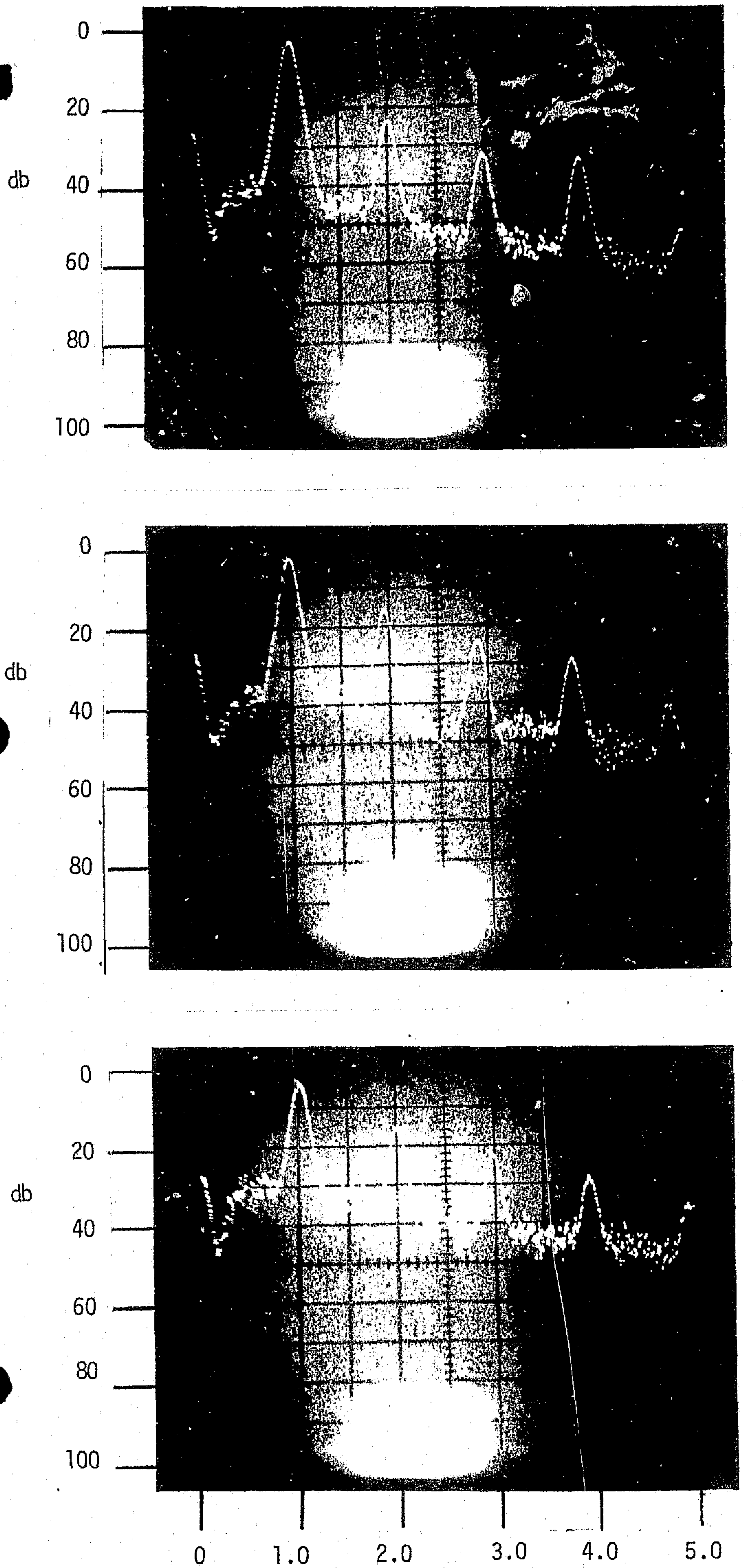


FIGURE 4-4
Db OUT vs FREQUENCY (TRANSMITTER TO RECEIVER)
PT-200 PORTABLE RADIO 10-28-76
ATTENUATION - 130, 144, 150 DB
5KHz PEAK DEVIATION
ATTENUATOR SETTING
-130 Db
-144 Db
-150 Db
0
-5
-10
-15
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
Audio

FIGURE 4-5

RECEIVER OUTPUT RESPONSE
1 kHz MODULATING TONE



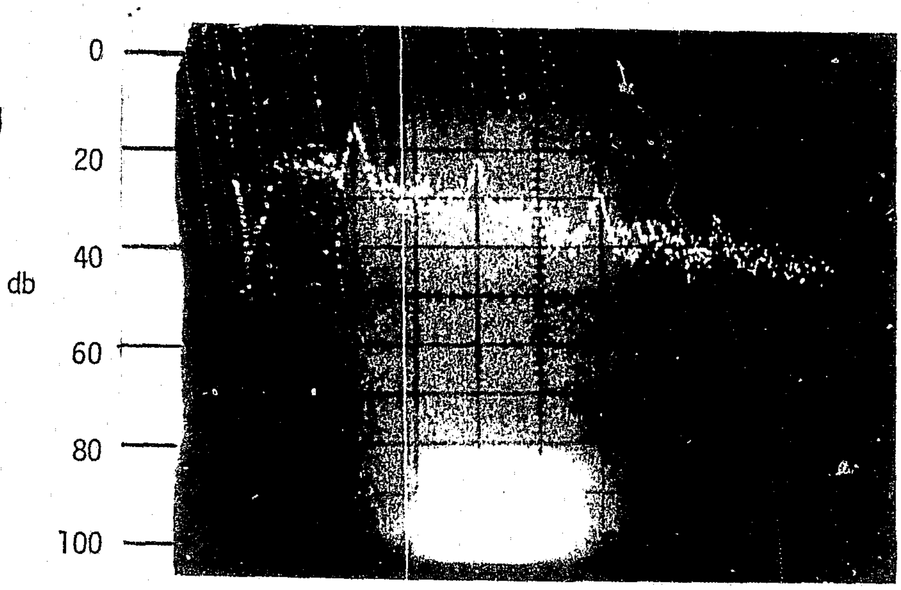
ATTENTUATOR SETTING
-100 Db
FREQUENCY 1KHz
TRANSMITTER TO RECIEVER

ATTENTUATOR SETTING
-120 Db
FREQUENCY 1KHz
TRANSMITTER TO RECIEVER

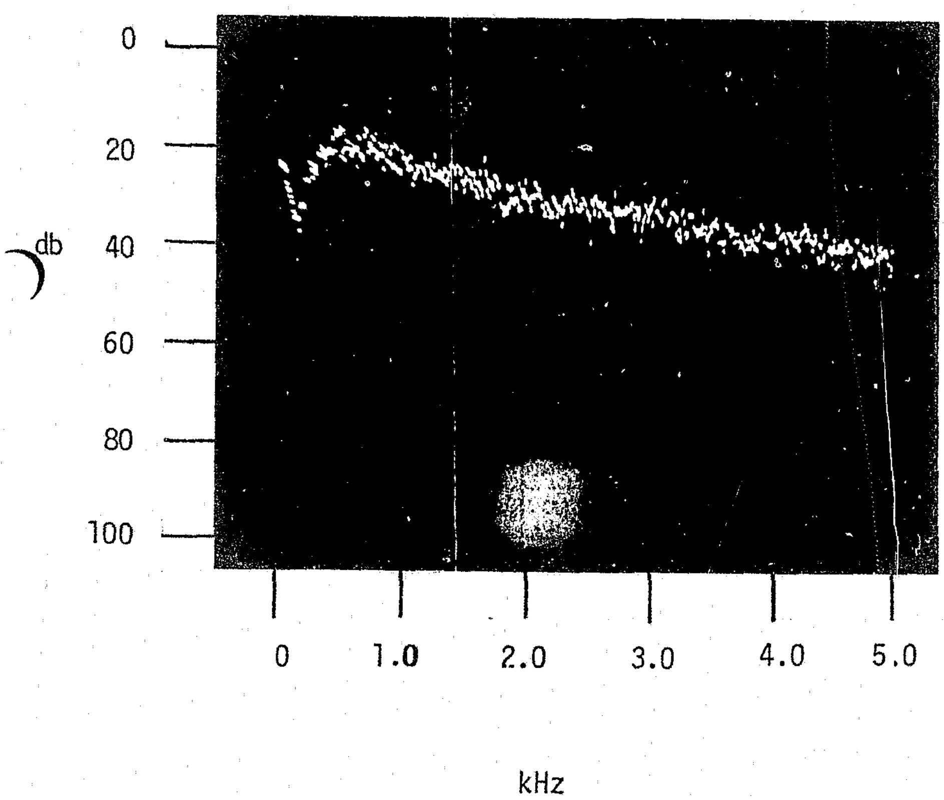
ATTENUATOR SETTING
-130 Db
FREQUENCY 1KHz
TRANSMITTER TO RECIEVER

FIGURE 4-5

(Continued)



ATTENTUATOR SETTING
-150 Db
FREQUENCY 1KHz
TRANSMITTER TO RECIEVER



ATTENTUATOR SETTING
-170 Db
FREQUENCY 1KHz
TRANSMITTER TO RECIEVER

response of the VDETS unit noted. The following qualitative results were obtained.

Distance	VDETS Response	Remarks
0.75 miles	Satisfactory	-
0.85	Satisfactory	-
1.4	Good	Retraining required
1.5	Marginal	Retraining required
2.0	Very poor	Voice signals unintelligible

It was concluded from the tests that VDETS units can be accessed by field radios with moderate success. The characteristic low signal-to-noise ratio environment, however, requires that all possible noise sources such as computer EMI be eliminated, and that the voice recognizer filtering algorithms be "tuned" to the characteristic signals of LMR networks. Although such an arrangement might be made to function, it is recommended that the recognizer units be placed in the field units.

Recommendations

The following recommendations are made based on the JPL tests:

- The frequency response characteristics of the LMR-VDETS combination should be considered in the design of the word detection process.
- The design of the VDETS should be matched to the receiver output noice spectrum and amptitude distribution.
- Equipment shielding should be considered.
- Transceiver microphone frequency response should be considered.

4.2 SCOPE ELECTRONIC, INC., TESTS*

The objective of this task was to test and evaluate the performance of voiced computer data entry and synthesized voice response over a standard police radio network using the Scope Electronics, Inc., (SEI) Voice Data Entry Terminal System (VDETS). Successful doemonstration of this capability would make it possible for field officers to make direct queries to data files without operator intervention, and to transmit field unit status and other call-for-service related information for computer processing and display to dispatcher personnel, or entry into a computer-aided dispatch system.

---

* This activity was sponsored under a separate LEAA contract with Scope Electronics, Inc., "Voice Recognition Over Radio." Contract No. 6-0815-J-LEAA See Final Progress Report, December 1976. The results of this task are included in the present report for convenience.

The task performed involved:

- Interfacing SEI's VDETS to SEI supplied transceivers operating at the same frequency as the District of Columbia radio network.
- Interfacing SEI's VDETS to the District of Columbia Metropolitan Police Department (DCMPD) radio network.
- Developing the VDETS software required for testing the accuracy and reliability of voiced computer data entry and synthesized voice response over the above standard police radio network.
- Testing and evaluating the accuracy and reliability of voiced computer data entry and synthesized voice response first at the SEI facility using the SEI supplied transceivers, then at the DCMPD facility using the DCMPD radio network.

Audio Interface

A block diagram of the VDETS system configuration is shown in Figure 4-6. The standard voice input device of the VDETS is a high quality noise canceling dynamic boom microphone. Therefore, it was necessary to modify the front end of the VDETS to accommodate the two-way radio networks. The appropriate impedance matching networks and voice in/voice out switching logic components were incorporated in the VDETS front end hardware and tested to verify proper operation. Since the function tones at 2250 and 2325 Hz were not present in the DCMPD radio networks, notch filtering of these tones was not necessary in the VDETS front end modifications. Also, a spectrum analysis of the 127.3 Hz squelch tone indicated that most of the tone energy was outside of the VDETS front end frequency band pass, such that it would not cause distortion of the voiced input frequency information. Consequently, no provisions were made in the front end for notching the squelch tone. Schematic diagrams of the SEI transceiver and DCMPD radio network interafaces are shown in Figures 4-7 and 4-8.

VDETS Software

The VDETS application software program uses the digits 0-9, the phonetic alphabet and the control words "READY", "ATTENTION", "COMPUTER", "RELAX", "CANCEL", "CORRECTION", and "OVER" as its dictionary. After user vocabulary training, the program requires the spoken entry of a 10 character alphanumeric data string. Upon recognition of the 10th character a synthesized voice verification transmission consisting of the 10 characters recognized by VDETS is provided to the user.

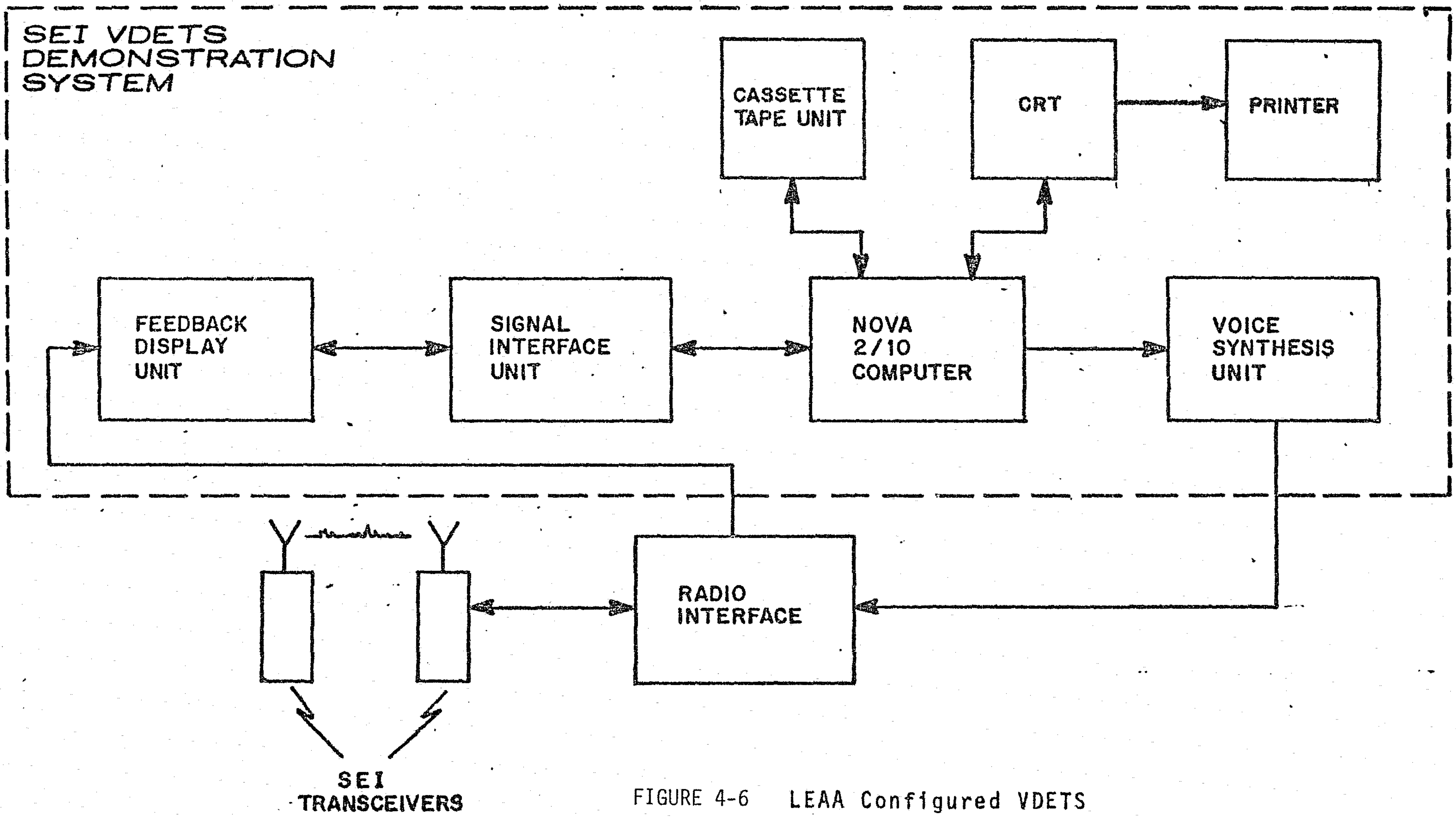


FIGURE 4-6 LEAA Configured VDETS

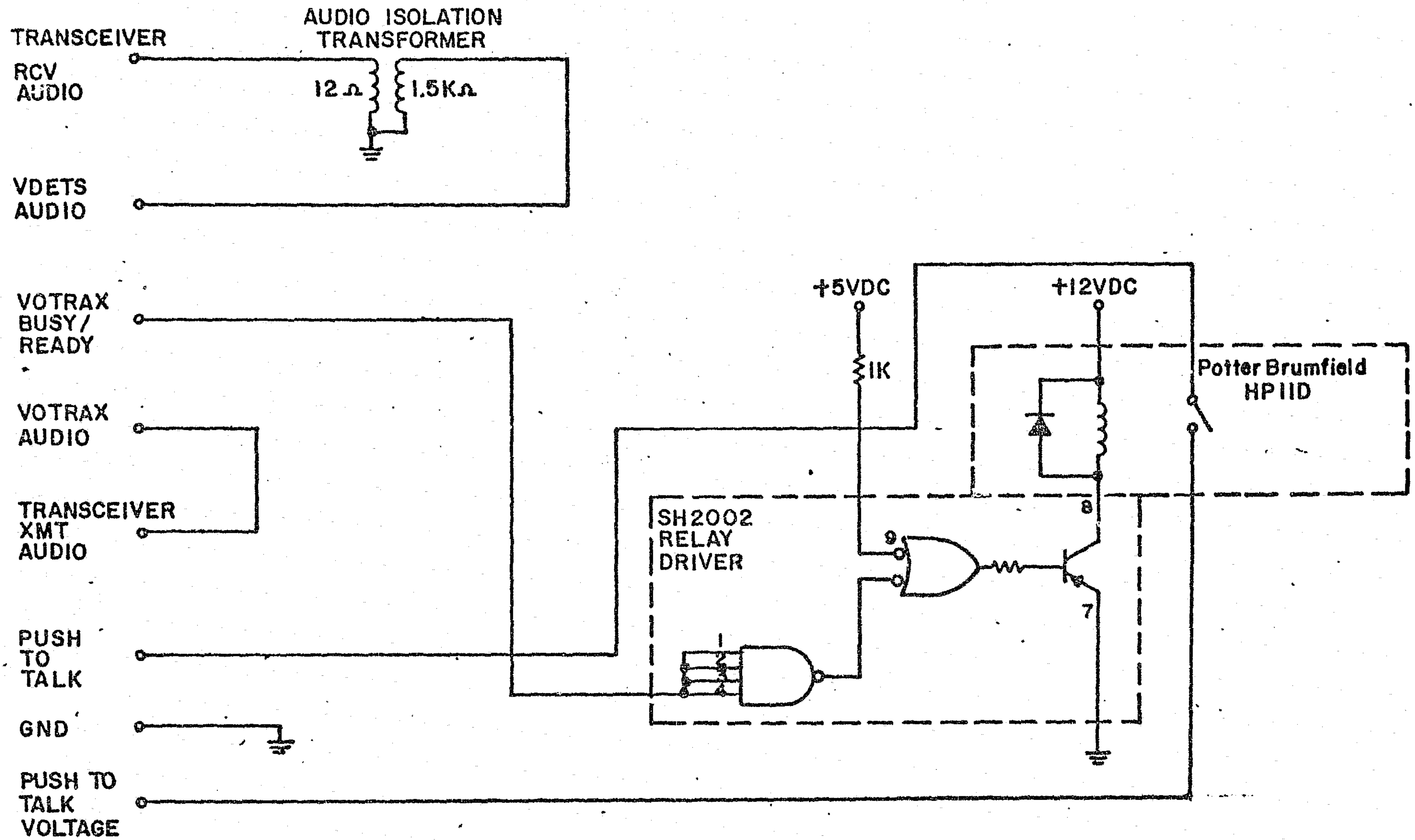


Figure 4-7 SEI Transceiver/VDETS Interface

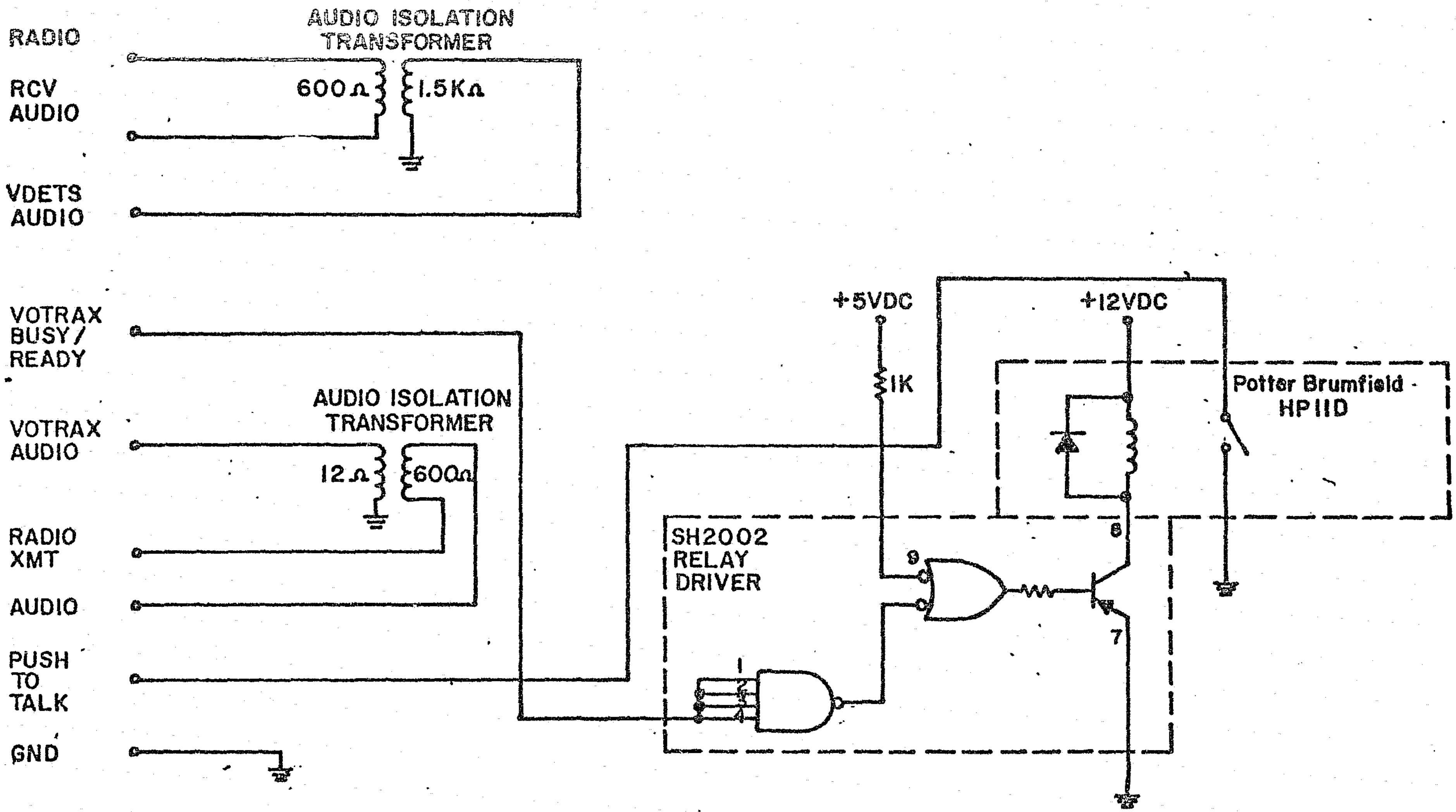


Figure 4-7 DCMPD Radio Network/VDETS Interface

Following the verification transmission the user indicates the correct recognition of the 10 character string by speaking the control word "OVER" or the incorrect recognition of the 10 character string by speaking the control word "CORRECTION". The control word "CORRECTION" allows for correcting recognition errors by repeating the 10 character string. In cases where difficulty arises in the attempt to correct a string or an incorrect string is spoken by the user, the control word "CANCEL" eliminates the 10 character string. The 10 character recognition sequence continues until terminated by the user speaking the control word "RELAX". This places the VDETS system in idle mode and disregards all further audio inputs except the control words "READY" "ATTENTION", "COMPUTER". The relax feature also provides for "locking out" from the VDETS all communications on the two-way radio network that are not pertinent to the specific function of the system. All of the 10 character strings spoken by the user including those recognized correctly, those recognized incorrectly and those cancelled by the user are recorded on the VDETS CRT and printer for evaluation and analysis.

Testing

Initial testing was conducted by several SEI users with the standard VDETS microphone input device to establish a bench mark for the two way radio testing. Ten or more alphanumeric character strings were spoken by each user. The recognition results were recorded and evaluated to verify the normal high recognition accuracy rate of the VDETS in the LEAA test configuration. Following the initial testing, the two way radio testing was conducted at the SEI facility using the SEI provided transceivers interfaced to the VDETS. The SEI users spoke ten or more alphanumeric character strings over the two-way radio interface at various distances from the receiver unit and the regognition accuracy results were recorded.

The VDETS was installed in the DCMPD headquarters and prepared for testing in the police radio network environment. During this time DCMPD and SEI personnel tested the VDETS in the police radio network environment using both mobile and portable radio equipment.

Test Results and Problems Encountered

A radio interference problem severely hampered the progress of the testing phase at the SEI facility. Because the transceivers of the VDETS operated at the DCMPD radio frequency, normal radio traffic interfered with the operation of the system during testing. A radio interference test verified that the SEI transmissions from the SEI facility in Reston, Virginia, were not received by the DCMPD radio network, while the DCMPD radio transmissions were received loud and clear by the SEI transceivers. Consequently, DCMPD radio transmissions frequently overpowered SEI radio transmissions resulting in uncertain user vocabulary training and test results. Only a limited amount of testing could be conducted at the SEI facility because of the interference. Transmissions from outside of the SEI building were virtually impossible. Most of the testing was ultimately conducted in an RF shielded room at SEI where outside radio transmission could not get through to the system. Therefore most of the testing was done from within 50 feet of the system. The limited valid testing at SEI indicated a good probability of success in testing at the DCMPD facility where superimposed radio transmissions would not be encountered. The results of the SEI facility testing are summarized in Table 4-2.

As anticipated the superimposed radio transmissions were not a problem at the DCMPD test facility. However, several days of testing resulted in unsuccessful operation of the system due to the highly variable signal-to-noise and spectral profile characteristics of the police radio network. These conditions created unreliable system training and test results. Therefore, after demonstrating the operation of the system, as mentioned above, the testing phase was concluded and the system was returened to SEI.

SEI did not identify or investigate the possible noise interference generated by the Nova minicomputer located in the VDETS unit. As noted in the JPL tests (see Section 4.1), Nova noise can seriously degrade VDETS performance when operating with low signal-to-noise ratio inputs, which are characteristic of mobile/portable radio networks. The improvements possible by properly shielding the equipments were not tested.

Table 4-2 SEI ACCURACY TESTS

INPUT DEVICE	USER	TEST NO.	NO. OF STRINGS	NO OF UTTERANCES	%CORRECT STRINGS	% CORRECT UTTERANCES
VDETS	A	1	39	195	95	99
Microphone		2	57	285	95	98
	B	1	20	220	90	99
Two-way	A	3	22	242	32	91
Radio		4	9	99	22	91
		5	17	187	65	96
		6	17	187	65	94
	B	2	12	131	50	94
		3	19	209	84	98
		4	19	209	89	99
		5	12	132	42	89

Note: 1. All two way radio testing was done from within 50 feet of the system except test B-5 which was done from the SEI facility parking lot, approximately 500 feet from the system.

2. All two-way radio testing encountered DCMPD radio inerference except when operated in the screen room.

5. CONCLUSIONS

It is concluded that voice recognition technology is currently available for application to law enforcement mobile communications. A good arrangement for this purpose is one in which the voice feature extractor is located in the patrol car and the supporting computer in the base station. In this way, consistently accurate input to the voice recognizer is assured, with minimal loading on the RF channel. A dispaly screen should be used in the patrol unit for convenience. A small, low-cost, low-power voice recognizer such as Speechlab has not been tested, but is illustrative of the equipment needed.

It is recommended that LEAA continue monitoring and testing the feasibility of in-car voice recognizer devices. Minimal funding is required for this purpose, in fact, an appropriate strategy for LEAA would be to encourage manufacturers to develop prototypes suitable for field testing. The ready technology and low costs of hardware modules should support such an approach.

Direct access to voice recognizers from mobile portable radios should be given lower priority at this time, although the technique can be made to function properly provided adequate precautions are taken to screen the equipments, whose performance is sensitive to the low signal to noise ratio environment.

It is recommended that LEAA sponsor a limited analysis effort to identify and develop concepts for networking communications systems that are comprised of voice recognizers located in patrol units supported by computer/classifier elements at the base station. Response times, channel loadings, and line protocol and control techniques should be investigated.

END