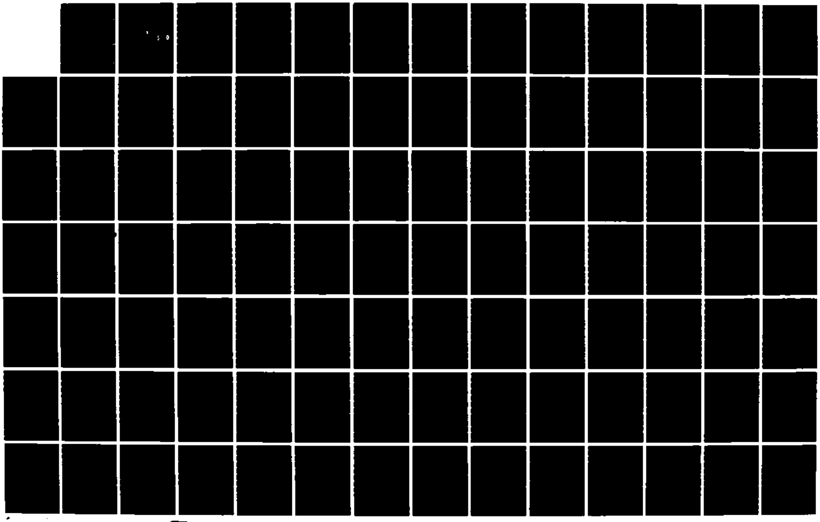
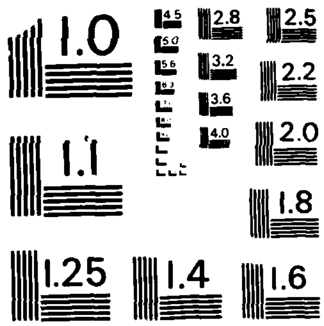
AD-A166 871 USING VOICE RECOGNITION AS AN INPUT MEDIUM TO THE JINTACCS (JOINT INTEROP..(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA E T HILL ET AL. MAR 86 1/1

UNCLASSIFIED F/G 17/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

USING VOICE RECOGNITION AS AN INPUT MEDIUM
TO THE JINTACCS AUTOMATED MESSAGE
PREPARATION SYSTEM (JAMPS)

by

Earl T. Hill
and
Leo B. Kotowski

March 1986

Thesis Advisor: Gary K. Poock

AD-A166871

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a SECURITY CLASSIFICATION AUTHORITY:

2b DECLASSIFICATION / DOWNGRADING SCHEDULE:

3 DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.

4 PERFORMING ORGANIZATION REPORT NUMBER(S):

5 MONITORING ORGANIZATION REPORT NUMBER(S):

6a. NAME OF PERFORMING ORGANIZATION: Naval Postgraduate School

6b OFFICE SYMBOL *(If applicable)*: 74

7a. NAME OF MONITORING ORGANIZATION: Naval Postgraduate School

6c. ADDRESS *(City, State, and ZIP Code)*: Monterey, California 93943-5000

7b. ADDRESS *(City, State, and ZIP Code)*: Monterey, California 93943-5000

8a NAME OF FUNDING / SPONSORING ORGANIZATION:

8b. OFFICE SYMBOL *(If applicable)*:

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER:

8c ADDRESS *(City, State, and ZIP Code)*:

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO

11 TITLE *(Include Security Classification)*: USING VOICE RECOGNITION AS AN INPUT MEDIUM TO THE JINTACCS AUTOMATED MESSAGE PREPARATION SYSTEM (JAMPS)

12 PERSONAL AUTHOR(S): Hill, Earl T. and Kotowski, Leo B.

13a TYPE OF REPORT: Master's Thesis

13b TIME COVERED: FROM ______ TO ______

14 DATE OF REPORT *(Year, Month, Day)*: 1986 March

15 PAGE COUNT: 96

16 SUPPLEMENTARY NOTATION:

17 COSATI CODES

FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*: Voice Recognition, Automatic Speech Recognition (ASR), JINTACCS, JAMPS, Message Preparation

19 ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This thesis investigates the interfacing of voice recognition, also known as automatic speech recognition (ASR), with the Joint Interoperability of Tactical Command and Control Systems (JINTACCS) Automated Message Preparation System (JAMPS). The voice recognition system we used is the Texas Instruments (TI) TI-SPEECH(tm) imbedded in the TI Portable Professional Computer (PPC). We were able to load the JAMPS software onto the TIPPC's hard disk. With the vocabulary we built, we ran the JAMPS software on the TIPPC using voice recognition. Our results indicate ASR has an application in message preparation during military operations. ASR could curtail the time to prepare messages, and thereby reduce the time element in the command and control process. We propose a measure of performance to test how much time might be saved by using ASR with JAMPS. We also suggest some areas for future research.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION: Unclassified

22a NAME OF RESPONSIBLE INDIVIDUAL: Gary K. Poock

22b TELEPHONE *(Include Area Code)*: 408-646-2636

22c OFFICE SYMBOL: 55Pk

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

Using Voice Recognition as an Input Medium to the JINTACCS Automated Message Preparation System (JAMPS)

by

Earl T. Hill
Captain, United States Army
B.S., United States Military Academy, 1977

and

Leo B. Kotowski
Captain, United States Air Force
B.S., United States Air Force Academy, 1980

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS TECHNOLOGY
(Command, Control and Communications)

from the

NAVAL POSTGRADUATE SCHOOL
March 1986

Authors: Earl T. Hill

Leo B. Kotowski

Approved by: Gary K. Poock, Thesis Advisor

Thomas J. Brown, Second Reader

Michael G. Sovereign, Chairman,
Joint Command, Control and Communications Academic Group

David A. Schrady,
Academic Dean

ABSTRACT

This thesis investigates the interfacing of voice recognition, also known as automatic speech recognition (ASR), with the Joint Interoperability of Tactical Command and Control Systems (JINTACCS) Automated Message Preparation System (JAMPS). The voice recognition system we used is the Texas Instruments (TI) TI-SPEECH(tm) imbedded in the TI Portable Professional Computer (PPC). We were able to load the JAMPS software onto the TIPPC's hard disk. With the vocabulary we built, we ran the JAMPS software on the TIPPC using voice recognition. Our results indicate ASR has an application in message preparation during military operations. ASR could curtail the time to prepare messages, and thereby reduce the time element in the command and control process. We propose a measure of performance to test how much time might be saved by using ASR with JAMPS. We also suggest some areas for future research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

We wish to express our gratitude to the dedicated people of the Tactical Air Forces Interoperability Group and of the MITRE Corporation who are working on the JINTACCS Automated Message Preparation System (JAMPS) at Langley Air Force Base, Virginia. Without their help and cooperation, this thesis would not have been possible. The following are the particular individuals we would like to thank:

<u>TAFIG</u>

Major Lester P. Tucker

<u>MITRE Corporation</u>

Paul R. Bigelow

Julio C. Fonseca

Robert W. Miller

John K. Thomasson

Special recognition goes to Jack Thomasson for all the extra hours he volunteered. He went out of his way to help us. The whole team effort by the hard-working MITRE personnel named above helped to open the door to some important future research into the integration of voice recognition with JAMPS. Their dedication highlights the motto they live by:

"YOU NEVER FAIL, UNTIL YOU STOP TRYING."

I. INTRODUCTION

In the last 5 to 10 years, the United States has increased the emphasis of the command and control (C2) process in military operations. For example, Dr. Joel S. Lawson, Sr., has proposed some informative models of the C2 process in his report, "The State Variables of a Command Control System." [Ref. 1: pp 93-96] If the U.S. military can SENSE the battlefield situation, PROCESS the sensed inputs to derive intelligence, COMPARE the processed information with the desired situation, DECIDE on a course of action, and then ACT more quickly than the enemy can do the SENSE-PROCESS-COMPARE-DECIDE-ACT process, then we can outfight the enemy. The idea is to have "effective" C2 so we can "beat the enemy to the punch." Within this C2 process, time is a critical element.

To get a common understanding before going further, let's look at the Department of Defense definition of command and control: [Ref. 2: p. 74]

> (C2) is the exercise of authority and direction by a properly designated commander over assigned forces in the accomplishment of the mission. Command and control functions are performed through an arrangement of personnel, equipment, communications, facilities, and procedures which are employed by a commander in planning, directing, coordinating, and controlling forces and operations in the accomplishment of the mission.

Looking at C2 in a more general view, it is the means by which a commander gets information about the situation which confronts him, how he reaches a decision in responding to the situation, and how his directions are communicated to his subordinates.

There are many projects designed to improve our C2 system. One such program is the Joint Interoperability of

Tactical Command and Control Systems (JINTACCS) Automated Message Preparation System (JAMPS). JINTACCS was established by the Joint Chiefs of Staff (JCS) so the military services would have common standards for communicating information during joint operations. JAMPS is the software to automate the preparation of these JINTACCS-standardized messages, thereby reducing some time elements in the C2 process. This thesis will look at the feasibility of improving the man-machine interaction by using voice recognition as an input medium for portions of JAMPS. Our specific goal is to integrate the voice recognition capability of the Texas Instruments Portable Professional Computer (equipped with TI-SPEECH(tm))[1] with the JAMPS software. We hope to relieve the operator from some of the mundane work associated with keyboard entry, focus his attention more on the message information, and thereby reduce the time element within the C2 process. In this paper we also suggest a measure of performance to test how much time might be saved by using voice recognition as an input method. Although the performance measure compares the manual method of preparing hard copy message traffic to the automated method (JAMPS), it can still be used to compare JAMPS to JAMPS with voice input. We conclude the thesis with our results and recommended areas for further research.

(Note: When we first started this research, we had hoped to interface the TI Portable Professional Computer and its imbedded TI-SPEECH card with the JAMPS II.V prototype hardware. However, toward the end of our thesis work, the JAMPS software had been upgraded so it could run on microcomputers such as the IBM PC-XT and IBM compatibles. This provided us with a different approach and an interesting result; the JAMPS software could run on the TI Portable

---

[1]TI-SPEECH is a trademark of Texas Instruments, Inc.

Professional Computer. This made the integration of voice recognition somewhat easier).

II. JINTACCS AND JAMPS

A. JINTACCS (JOINT INTEROPERABILITY OF TACTICAL C2 SYSTEMS)

The basic goal of the JINTACCS Program is to develop plans to achieve compatibility and interoperability of the military's command and control (C2) systems in joint air, ground, and amphibious operations [Ref. 3: p. 1]. Specifically, the JINTACCS Program Office in Washington, D. C., is responsible for developing standardized message formats and procedural interfaces between tactical C2 systems in joint/combined operations [Ref. 4: p. 37]. The program was a necessary response to the problems the U.S. military faced when trying to operate together with the various tactical C2 elements having their own equipment, procedures, and standards.

JINTACCS had its origins back in April, 1971, when the JCS established the program Joint Interoperability for Tactical C2 Systems in Support of Ground and Amphibious Military Operations (GAMO for short) [Ref. 5: p. 2]. GAMO's objective was to ensure interoperability and operational effectiveness of tactical C2 systems used to support ground and amphibious operations. The GAMO Program discussed the development of character oriented messages (COM) standards. COMs are messages which are human readable, yet can be processed by machines. A new acronym used in place of COMs is message text formats (MTFs). In this paper, COMs will be a generic reference to messages which people can read, while MTFs will pertain strictly to JINTACCS message standards/ formats.

GAMO also picked up the responsibility for developing bit oriented messages (BOM) standards [Ref. 5: p. 2]. BOMs are ideal for machine processing, but require some form of translation for people to read. BOM standards are commonly known as tactical digital information link (TADIL) standards. Since JAMPS does not deal with BOMs, no more will be said in this thesis about BOMs.

With the issuance of JCS Memorandum SM-184-78 in 1978, the JCS reorganized the GAMO Program [Ref. 5: p. 4]. The program name was changed to JINTACCS. Although the program's objectives remained essentially the same, the JINTACCS program now had more authority in order to achieve the interoperability of tactical C2 systems. JINTACCS does not specify message standards for each service's internal procedures, but rather the message procedures and standards of services involved in joint operations [Ref. 6: p. 56]. With its reorganization and tasking for proven results by 1985, the JINTACCS program set up a plan to develop, test, and demonstrate these message formats and procedural interfaces.

The most recent change for the JINTACCS Program Office occurred 5 July 1984. This is when the Department of Defense formally established the Joint Tactical Command, Control and Communications Agency (JTC3A) at Fort Monmouth, New Jersey. The JTC3A's mission is to [Ref. 7: foreword]

> ensure the interoperability of tactical . . . (C3) systems for joint or combined operations . . . and accomplishes this through the development and maintenance of a joint architecture, interface standards, and interface definitions for tactical and mobile C3 systems.

With its formation, the JTC3A now had responsibility for the JINTACCS Program as well as many other C2 related programs. Yet the JINTACCS objectives remained the same.

So for a quick recap and without regard to the BOM (i.e., TADIL) standards, the specific purposes of the JINTACCS MTFs are [Ref. 8: p. 1-3]

- provide same message formats for all services in order to exchange tactical information between operational facilities;
- provide messages which computers can process;
- provide messages which people can read.

1. <u>Message Text Formats</u>

COMs are record message traffic. The usual procedure to send COMs is for someone to type on a DD Form 173 (Standard Message Form) or to type or handwrite on a Joint Message Form, then bring it to a communications center for processing and transmission. (This paper only references the Joint Message Form). As mentioned earlier, the JINTACCS requirement is for MTFs to be machine processed <u>and</u> human readable. This need is especially important when considering fully automated C2 systems. The MTFs will be read by people and will also be used to automatically update data bases or reformat messages for further transmission. JINTACCS message standards now encompass strictly formatted messages for five JINTACCS functional segments: intelligence, air operations, fire support, maritime operations, and operations control [Ref. 3: p. 7]. Presently there are about 135 standardized messages for these 5 segments. At the time of this writing, the JTC3A's JINTACCS Program Office was making plans to include a sixth functional segment, combat service support [Ref. 7: pp. 5-6]. Implementation of this segment is expected in 1989 and should add about 60 more MTFs. JTC3A is also incorporating a seventh segment for special operations [Ref. 9: p. 53].

2. <u>Message Composition</u>

JINTACCS messages, like all messages, have a heading, text, and ending, as shown in Figure 2.1 [Ref. 8: p. 1-3]. The heading contains the addressing and routing information for the message. The text is the body of the message. Both the heading and ending formats depend on the communications system being used.

The text of JINTACCS messages consists of two basic elements: the set and the field. Fields make up sets. There are also several categories of sets, which are shown in Figure 2.2 [Ref. 8: pp. 1-4 to 1-5].

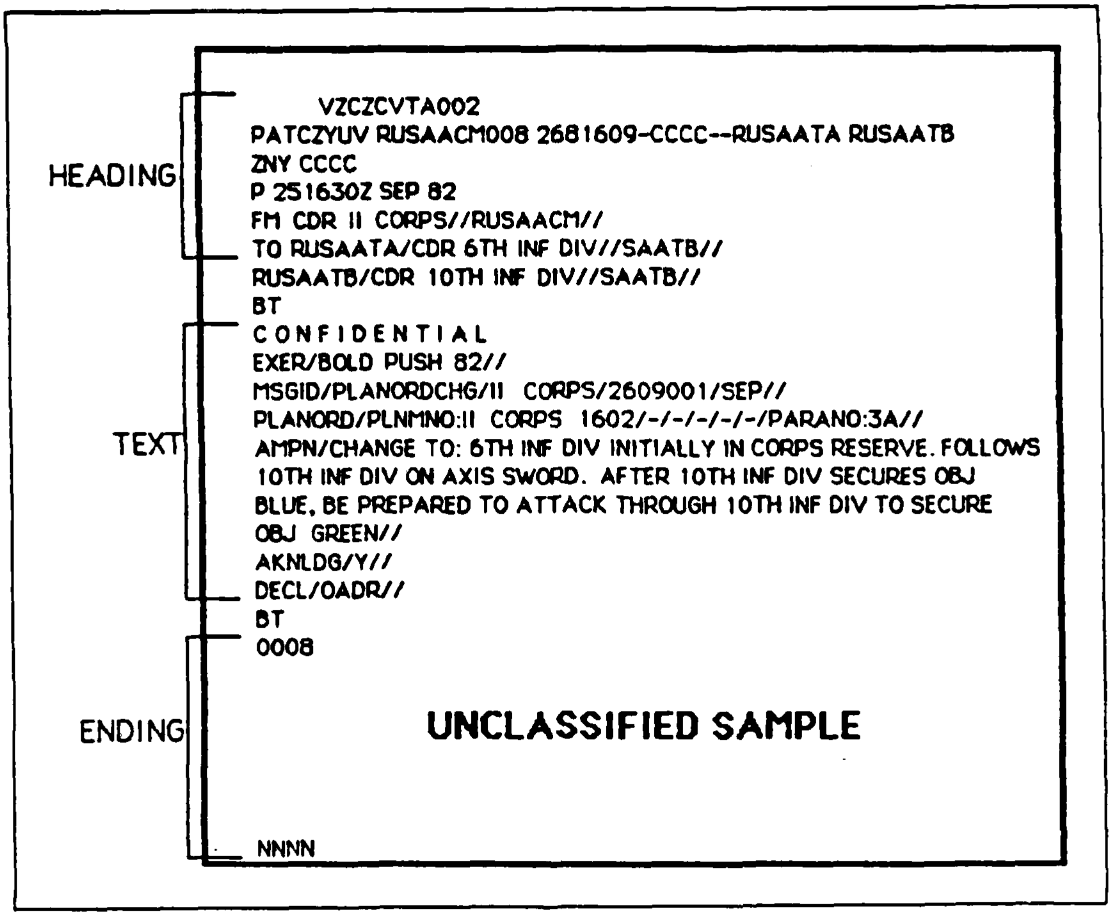


Figure 2.1 Message Components

- SET is the basic building block of all JINTACCS messages. A set is similar to a sentence and has information about one subject.
- FIELD is the basic building block of the set. You can think of the field as a word within the sentence, the set.
- LINEAR SET is a group of fields in a line across the page.
- COLUMNAR SET is a group of fields arranged in columns (i.e., tabular form).
- FREE TEXT SET is a set entered in lines, columns, sentences, or paragraphs. Free text sets provide information which does not fit into other types of sets.
- SET/FIELD NAME is a name to help you recognize the set or field. The name tells you what is in the set/field. In columnar sets, the name is sometimes referred to as columnar headers.

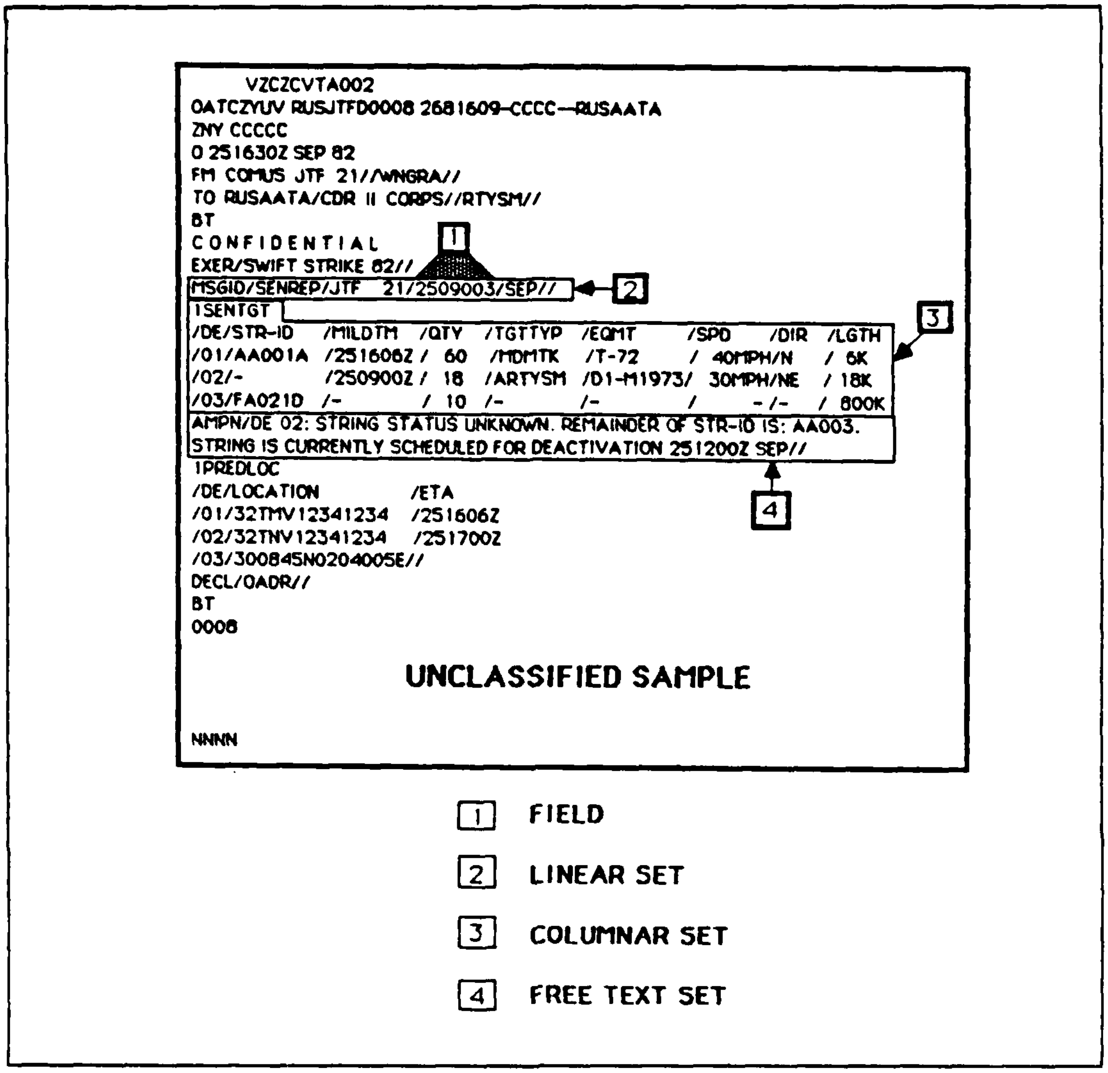


Figure 2.2 Message Text Components

Sets and fields have usage categories (such as mandatory, conditional, optional, and repeatable) which tell the JAMPS operator if the set or field must or can be used in a message, or if it can be used more than once. This information will appear at the bottom of the video display as the system user is filling out message templates. Conditional means the operator does or does not have to complete a field depending on whether he did or did not complete a previous field.

3. Formatting Rules for JINTACCS Messages

There are general rules which govern the input of information into the fields and sets of JINTACCS messages. [Ref. 8: pp. 2-3 to 2-5]

- The allowable characters for input are
  - all capital letters, A-Z;
  - all numbers, 0-9;
  - blank spaces;
  - special characters . , : ( ) ? - /.
- The slant bar (/) is a beginning-of-field marker and should not be used for anything else in linear or columnar sets.
- The double slant bar (//) indicates the end of a field. It acts as a period in a sentence. The slant bars must be on the same line.
- Do not split a field between lines.
- You cannot enter more than 69 characters on a line.
- The hyphen (-) is the no data sign. You would use this when you must make some entry into a field or set, but you do not have the information. There are specific rules guiding the entries into mandatory sets and fields, and conditional and optional fields.
- The special characters which are NOT ALLOWED IN JINTACCS MESSAGES are ! ; " @ # % $ & * + = and the degrees symbol. Operators should take special care not to include these symbols in free text sets.

4. Adopting JINTACCS Standards

A series of tests and demonstrations of the new message standards have taken place from 1981 through 1985. SOLID SHIELD '85 was the last operational effectiveness demonstration (OED) of the MTF standards before implementation. The JCS has mandated that the JINTACCS-developed MTF standards for the first five functional segments will be implemented by 30 September 1986. The services will use them in any joint service operation thereafter.

B. JAMPS OVERVIEW

Early testing (to include the 1981 JINTACCS OED) of formatted messages using JINTACCS standards showed some form of automation would be necessary [Ref. 10: p. 2-1]. The messages were difficult to manually prepare and often full of errors. This problem led to timeliness requirements not being met for messages. Obviously word processing and automatic data processing could provide a solution. This is when the Air Force, along with the MITRE Corporation, developed the JINTACCS Automated Message Preparation System (JAMPS). The Air Force organizations included the Tactical Air Forces Interoperability Group (TAFIG) at Langley Air Force Base, Virginia, and Electronic Systems Division at Hanscom Air Force Base, Massachusetts.

In developing the software, the JAMPS effort explored message preparation techniques, software/hardware interfacing, configuration management, and the capabilities of JINTACCS standards [Ref. 6: p. 57]. The software uses data tables to select, prepare, store and retrieve, review, transmit, and receive JINTACCS-formatted messages (i. e., MTFs). The JAMPS software is designed to make message rules and formats as transparent as possible to the operators during the operational effectiveness demonstrations [Ref. 6: p. 57]. This is a crucial design feature considering the complexity of the JINTACCS standards.

1. JAMPS Beginning

The origin of JAMPS coincided with a MITRE Corporation project to support the JINTACCS Program. MITRE had created the message encoding/decoding technique called Interoperability Through Structured Message Exchange (ITSME) [Ref. 3: p. 6]. MITRE upgraded ITSME and gave it to the Air Force as JAMPS, later to become JAMPS I. JAMPS I software is written in "C" programming language under the UNIX operating system and runs on the PDP 11/70 mainframe computer. JAMPS I also uses Perkins-Elmer OWL 1200 display terminals. After many enhancements and installation of the software onto portable hardware, JAMPS I became JAMPS II.

2. JAMPS II Hardware

The new hardware for JAMPS II was based on the Digital Equipment Corporation (DEC) LSI-11/23 microprocessor with 256 kilobytes of random access memory (RAM) [Ref. 3: p. 27]. Other hardware features include [Ref. 11: pp. 12-13]

- a Data Systems Design (DSD) 880/20 combination disk drive with a 20 megabyte Winchester hard disk and a 1 megabyte, 8 inch floppy disk drive;
- a DEC VT103 video terminal with keyboard, display, and a card cage with the microprocessor, memory and peripheral cards;
- a DEC MSV11-LK 256 kilobyte RAM board;
- a DEC DLV11-J 4-port RS-232 multiplexer;
- a DEC DLV11-E asynchronous line interface;
- a DSD 8832 interface board;
- three empty dual-connector card slots with Q-bus interface;
- a Centronics 150-3 dot matrix printer;
- a bus interface unit (BIU), the Digital Communications Corporation (DCC) BIU Model 4024.

3. JAMPS II.V Hardware

The latest prototype, JAMPS II.V, is a further upgrade from the previous JAMPS workstations. JAMPS is still written in "C" language and runs under the UNIX operating system. The upgrading replaced the JAMPS II microprocessor, RAM, and disk drives. JAMPS II.V now consists of a DEC VT103 video terminal and a Cambridge Digital System (CDS) 94 system.[2] This CDS 94 consists of

- a chassis (SIGMA SA-H102) with Q-bus backplane, 8 Quad slots, 22 bit addressing;
- a Cadmus 9000 MC68010 microprocessor board;

---

[2] "JAMPS II.V Operator's Manual," Appendix G, p. 1. This list is an update to the present manual's hardware list found in Appendix G.

- a PCS QS512 512 kilobyte memory board (for RAM), which is dual-ported and expandable;
- a DEC DLV11-J 4-port RS-232 interface board;
- a DEC DLV11-E single-port full modem control board;
- a DEC DZV11-M 4-port multiplexer board;
- an Emulex SC02/C Winchester controller board;
- a SIGMA SDC-RXV31 floppy controller board;
- a Fujitsu M2312 84 megabyte Winchester drive (hard disk);
- a Tandon TM848 floppy drive (8 inch floppy disk);
- Winchester disk bootstrap software;
- diagnostic firmware.

In addition to the above components, the JAMPS II.V workstation has a DCC BIU Model 4024, and a Datasec DSP-225 ink jet printer. The benefit of the JAMPS workstation upgrade from version II to II.V is a much improved response time.

(Note: JAMPS II and II.V imply both hardware and software, whereas JAMPS, by itself, implies just the software. In a generic sense, JAMPS also refers to the concept of automated message preparation).

4. JAMPS Software on Microcomputers

As of the date of this writing, TAFIG and MITRE had just finished upgrading JAMPS to run on microcomputers. JAMPS software can now run on International Business Machines (IBM) PC-XT microcomputers and IBM compatibles, such as the Zenith Z-150. This seems a logical and necessary path to follow since desktop computers are in widespread use in the military.

5. Features and Capabilities

JAMPS provides a variety of menus for available messages. The main menu provides the categories seen in Table 1 [Ref. 10: p. 8-5]. (Note that combat service support and special operations do not appear). Within categories 2 through 6 there are over 130 message formats within the data tables. Upon choosing the message type you want to send, you would be presented with a mask mode template. This is a fill-in-the-blanks template of up to 20 lines. Operators enter information into empty data fields by using the keyboard. Figure 2.3 is the mask mode template for an air base status (ABSTAT) report [Ref. 3: p. 13].

TABLE 1

JAMPS MAIN MENU

1	ADMIN
2	AIR OPERATIONS
3	OPS CONTROL
4	INTEL
5	FIRE SUPPORT
6	MARITIME
7	Messages Saved
8	Messages Transmitted
9	Messages Received
10	MERGE FILES
11	*** Quit ***

Should the operator want, JAMPS will display "map characters" in the empty data fields where the cursor is located or "tabbed" to. The map characters show the allowable data types (such as numeric, alphabetic, blank, and special symbols) which go into the data field. The operator cannot proceed in filling out the message if he enters the wrong data types (i.e., entering a number in an alphabetic map character position).

A HELP function is also available to assist the operator in filling a field with data. The HELP function

```
MSGID/MSGTYPE: ABSTAT                  /ORIGINATR: ____________________
/SERNO: _______ /MNTHNM: ___ /QLF: ___ /NO: ___ //
REF/SRLIDTFR: _ /MSGTYPE: ____________________
/ORIGINATR: ____________________ /MILDAYTI: _______ /SERNO: _______
/SPECNOTA: _____ /NASISCOD: ___ //
AMPN/ ______________________________________________________________ //
NARR/ ______________________________________________________________ //
CANX/MSGTYPE: ____________________ /ORIGINATR: ____________________
/MILDAYTI: _______ /SERNO: _______ /SPECNOTA: _____ /NASISCOD: ___ //
BASESTAT/DE: __ /BASE: ____________________ /ICAO: ______ /OPS: aaa
/EFFDAT: _________ /ETRO: _________ /ABLOC: _______________ /ELEV: _____ //
AMPN/ ______________________________________________________________ //
BRUNWAY
/DE RWYDSG  OPS TIME       TMEDES LGTH WDTH RWYCOMP WTCAP  LTG LTG LTG
 __ _______ ___ _________ ______  ___  ___ _______ ______ __  __  __
 __ _______ ___ _________ ______  ___  ___ _______ ______ __  __  __ //
AMPN/ ______________________________________________________________ //
BARREST
/DE RWYDSG ARSYS POS  OPS TIME       TMEDES CMNT
 __ ___     _____ ____ ___ _________ ______ _______________________
```

Figure 2.3 B704 Mask Mode Template

displays, for the field the cursor is in, a screen of information concerning that field. The operator can fill out an entire message using just the HELP displays by entering the "conversation mode." [Ref. 12: p. 5] In conversation mode, fields are presented one at a time as part of the HELP function. When the operator finishes with the current field and pushes the TAB key, the next field is shown in a HELP display. At any time the operator may revert to mask mode. The operator can edit the message very easily since the JAMPS software has many of the common word processing features. Editing includes calling in additional data fields. These modified message formats can then be stored for later recall.

The operator can review a message he is preparing by entering the "review mode." Review mode deletes unused data fields and trailing blanks. This also happens when the operator transmits a message. While in review mode, the system user may hit a VALIDATE key to make JAMPS check the message for data-sequencing errors. These types of errors are violations of certain JINTACCS formatting rules and occur when the user has failed to complete a mandatory data field or set.

As depicted in Figure 2.4, information about JINTACCS message formats resides in data tables external to the program code [Refs. 3,6: pp. 19, 57]. This provides two advantages. First, the cost of changing message formats will be reduced. Second, other C2 software programs can access the same data tables. This capability will make the remote updating of JINTACCS standards much easier. These external data tables will give tactical commanders and communicators the ability to deal with information exchange needs not originally foreseen by JINTACCS architects [Ref. 10: p. 2-2]. Figure 2.5 outlines the JAMPS software architecture [Ref. 3: p. 21].

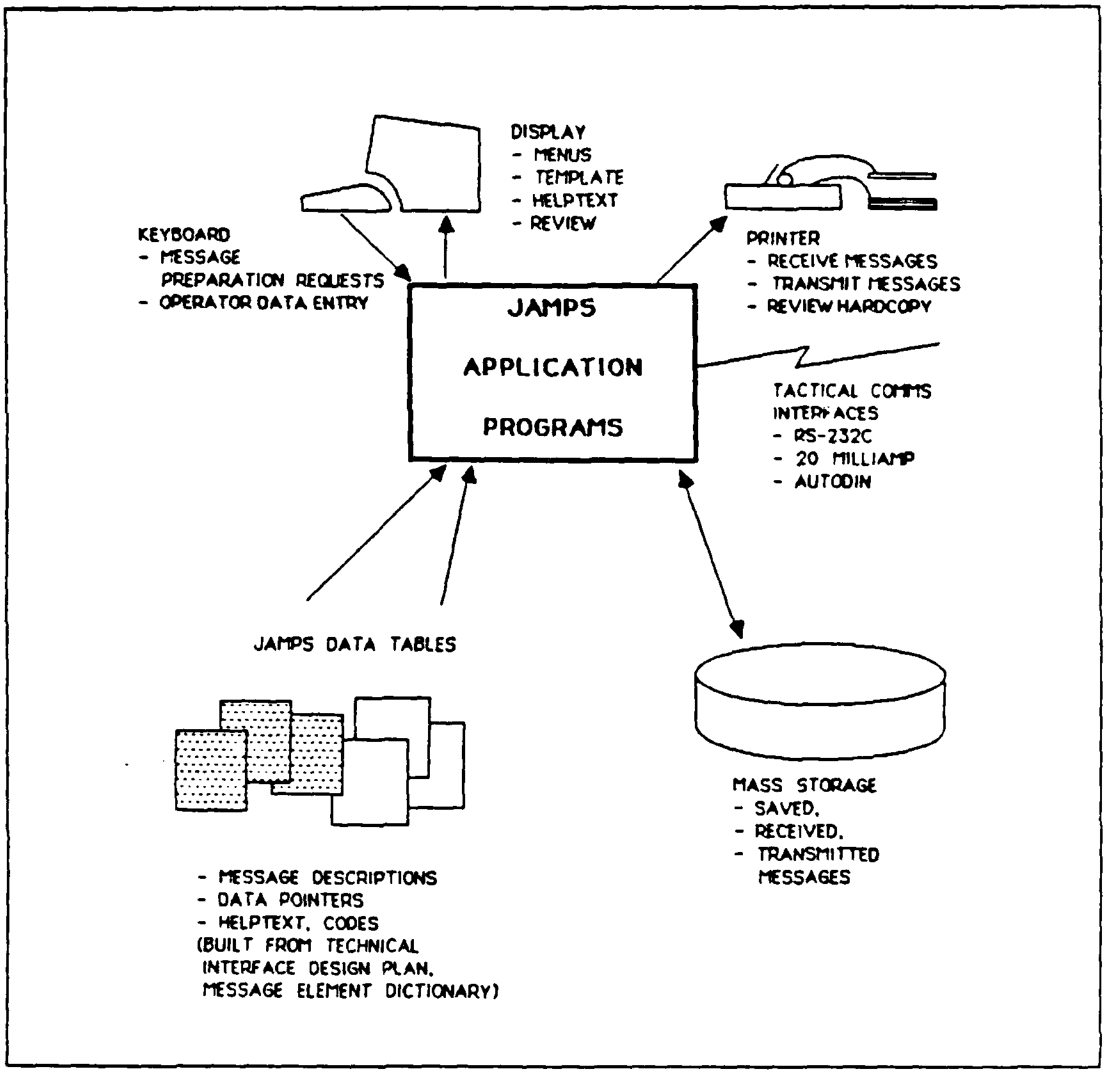


Figure 2.4 Table Driven Software

JAMPS is compatible with an encode/decode scheme which will let dissimilar systems exchange and process JINTACCS messages, even if these systems are tuned to different message versions of JINTACCS [Ref. 10: p. 3-1]. This is true in Europe, where some teletype messages may take more than 12 hours to reach their destination because of the number of manual store-and-forward switches. Even if all C2 facilities updated message standards at the same

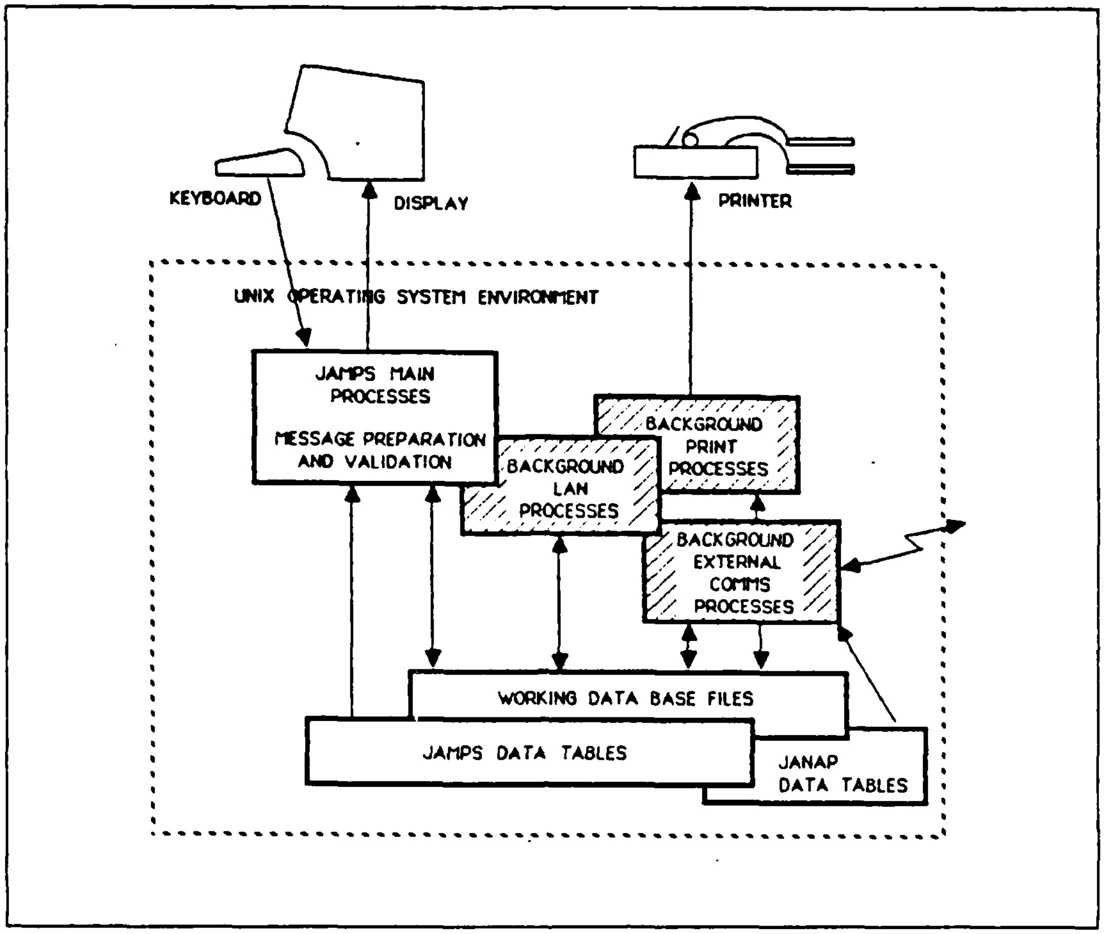


Figure 2.5 JAMPS Software Architecture

time, there would still be some lag time in which two different versions of standards were in effect. JAMPS could accomodate these multiple editions during the transition. The encode/decode methodology further expedites message transfer by compressing (or "packing") the message 30 to 50 percent before transmission, and "unpacking" it at the receiving end.

There are actually two sets of external data tables. One set is the master set, while the other set, a draft set, is a machine-efficient derivative [Ref. 10: p. 3-1]. Should the draft files be accidentally destroyed, then JAMPS automatically regenerates new draft files.

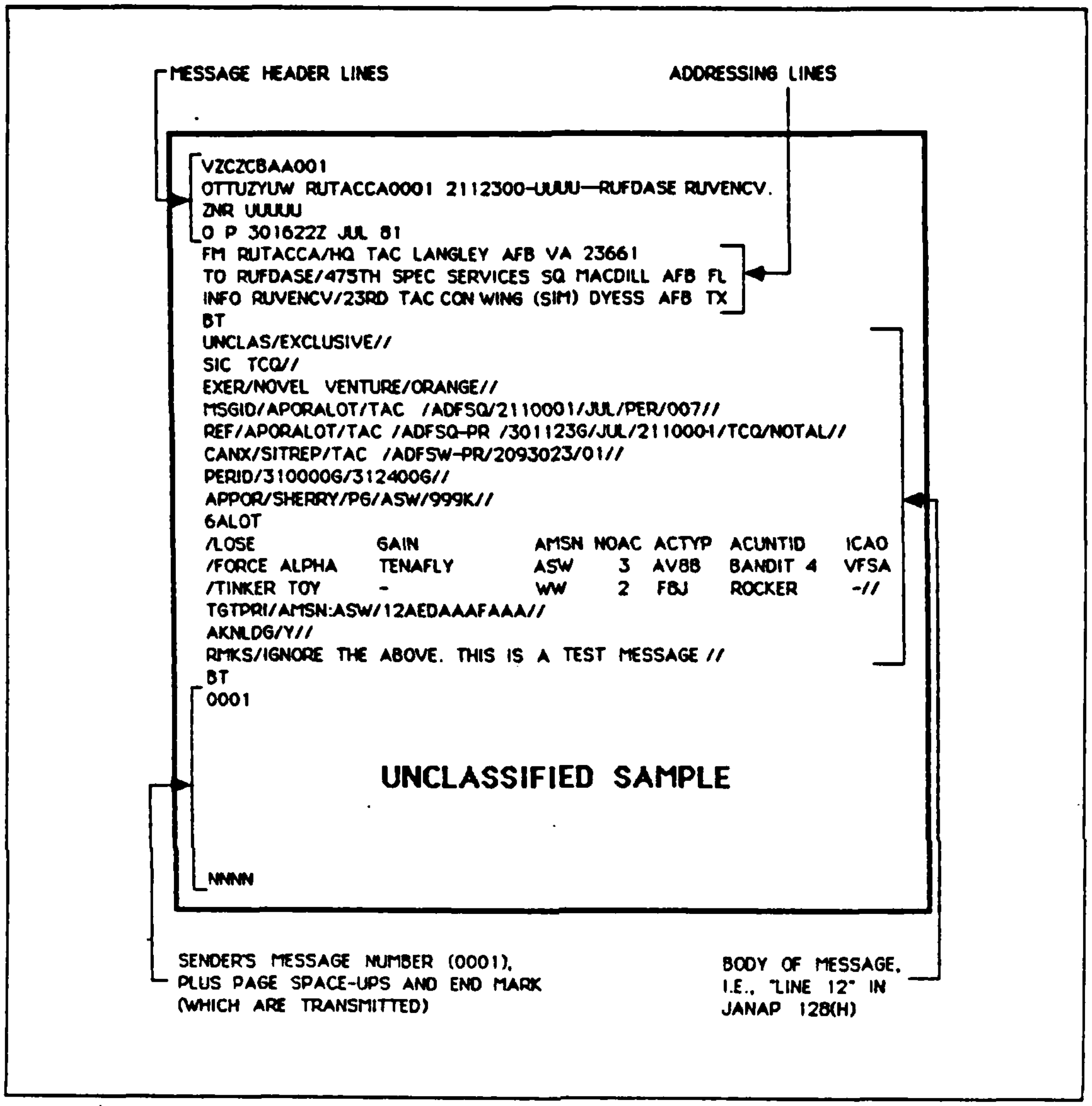


Figure 2.6 Message Printed in Transmission Format

The JAMPS software needs message header and trailer information to transmit the JINTACCS message. See Figure 2.6 for an example of a message in transmission format [Ref. 12: p. 32]. The message header contains addressing and routing information. Separate addressing lines also contain FROM, TO, and INFO addressees. This information must conform to the rules in the Joint Army, Navy, Air Force Publication

(JANAP) Automated Digital Network (AUTODIN) Operating Procedures (JANAP-128(H)) to allow transmission over military communications systems. A JINTACCS message actually becomes the main body of a JANAP 128(H) message. In particular it will be in AUTODIN PLAINDRESS message form whereby, the addressing information is sent in the clear, even if the body of the message is encrypted [Ref. 12: p. 4]. Like the JINTACCS data tables, there are JANAP data tables which the applications program can access for the addressing and routing information. Once the operator hits the transmit key for the JINTACCS message, JANAP 128(H) headers and trailers are added and the message then goes into the AUTODIN circuits [Ref. 11: p. 28].

JAMPS interfaces with the AUTODIN communications system by AUTODIN Mode I or Mode II. Mode I is a highly automated protocol which does the interfacing between AUTODIN and the JAMPS workstation. Mode II is a manual interface needing much more operator involvement. Finally, after the JINTACCS message receives the JANAP header and trailer information from the stored tables, the message is printed out on a line printer, converted from American Standard Code for Information Interchange (ASCII) to 5-level Baudot, then transmitted over external communications links [Ref. 11: p. 28]. JAMPS automatically keeps a copy of the transmitted message for future reference, retransmission, or change.

When the JAMPS workstation receives a message, the message is automatically printed out and stored; and the JAMPS workstation alerts the operator to the message's arrival [Ref. 11: p. 15]. Once the JAMPS system has stored the message, then the operator may review, reprint, or delete the message. The received message, like a transmitted message, is in 5-level Baudot code. The Baudot code is then converted to ASCII.

JAMPS workstations are compatible with the AN/TYC-39 message switch. The AN/TYC-39 is a 25 or 50-line tactical, automatic store-and-forward message switch for record (hard copy) communications [Ref. 13: pp. 448-449]. This message switch directs the message traffic in its tactical network. It features automatic access and tandem message processing, automatic message protection and management, and traffic segregation. The AN/TYC-39 message switch is already operational in Europe.

JAMPS II.V workstations can be configured for a variety of operations, depending on the tactical C2 environment. The various configurations are listed below. [Ref. 10: pp. 3-2 to 3-3]

- Single workstation not connected to a communications link. Draft messages, when approved for release, would be hand-carried to a communications center for transmission.
- Multiple workstations not linked together or to a communications link. Draft messages, once approved for release, would be hand-carried to a communications center for transmission.
- Single workstation connected to a communications link.
- Multiple workstations with one station connected to a communications link. Workstations would produce draft messages and save a copy on a diskette. If the message is approved for release, the diskette will be brought to the workstation connected to the communications gateway for transmission.
- Multiple workstations networked together with one station connected to a communications link. Workstations can forward "mail," messages, or portions of messages to the gateway workstation. This local area network can accomodate up to three computers with one acting as a central node and the point of entry to the communications channels. Each of the three computers can host up to six workstations.

As you can see from these various configurations, there is much flexibility and responsiveness in using the JAMPS II.V. The workstation arrangements would depend on hardware availability, commander's needs, and the tactical C2 environment.

C. JAMPS IN REVIEW

JAMPS is a system of hardware and software designed for military personnel in sending tactical communications during joint operations. The JINTACCS Program has explored the use of common message standards for the military services in joint tactical operations. The JINTACCS Automated Message Preparation System has tested the concept of automating the more than 130 JINTACCS messages. JAMPS II and II.V are portable computers with the JAMPS software and the necessary input/output and storage devices. Just a few of the main features that help the operator prepare messages are fill-in-the-blank message templates; HELP menus; and message review, storage and retrieval. Military forces can link together JAMPS workstations in a local area network and interface with external military communications systems. The JAMPS developers designed the system with people in mind so they made JINTACCS message rules and formats as transparent as possible to the system user.

The next chapter will look at voice recognition in general. The fourth chapter will explain the TI-SPEECH ASR system.

III. SPEECH RECOGNITION

As stated in a 1984 article by George M. White, speech recognition is definitely an idea whose time is coming [Ref. 14: p. 213]. Speaking to one's computer may seem an unusual concept; but, as automatic speech recognition (ASR) technology improves and becomes more affordable, it will be as natural a method of input to a computer as the keyboard is now. ASR is generally thought to be the most difficult and complex problem to solve in the area of voice processing. Other aspects of voice processing, including voice synthesis, analysis, and encryption, are more easily defined and treated. ASR includes aspects of these fields and is therefore far more difficult to bound and solve its problems. Many large companies, universities, and the Department of Defense have devoted a great deal of energy to ASR research. The progress has not been as rapid as hoped for, but it has been steady. Because of the wide range of disciplines involved and the time already spent on the challenges of ASR, progress should be considered evolutionary, not revolutionary [Ref. 14: p. 213].

This chapter will provide general background on ASR technology as it now exists. Brief discussions will address the differences between discrete and continuous speech recognition systems and between speaker dependent and independent systems. One section will cover some of the voice recognition systems available at the Naval Postgraduate School. Another section will review some pertinent studies also done at this school. This information provides a list of references should someone wish to do further studies related to this thesis.

A. TYPES OF SPEECH RECOGNITION SYSTEMS

Studies have shown that the development of a voice-interactive system, as seen in Figure 3.1(a) [Ref. 15: p. 37], must consider the entire system and not just the speech component.

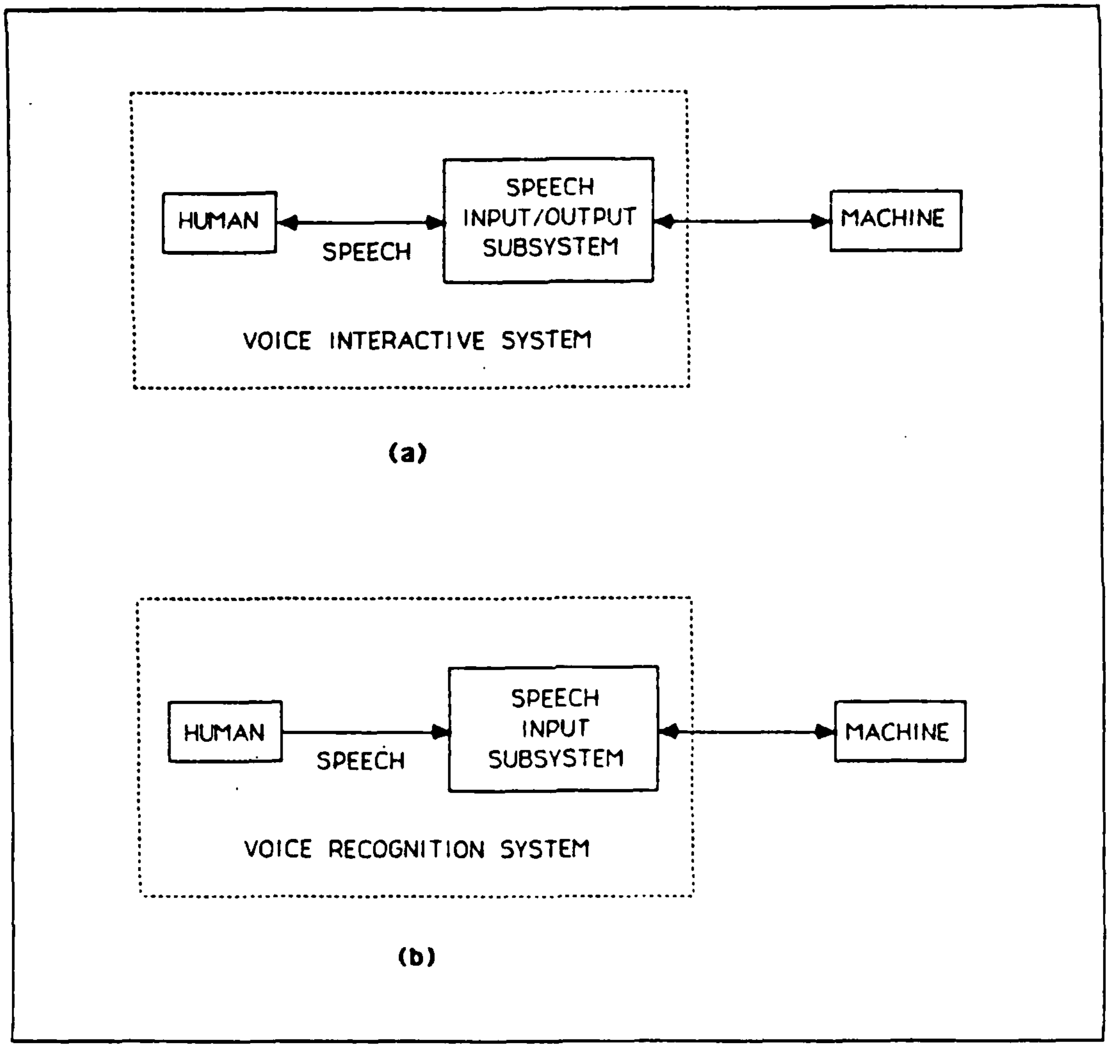


Figure 3.1 Voice Interactive and Recognition Systems

The voice interactive system is a man-machine interaction between a person and a speech input/output (I/O) subsystem. The I/O subsystem includes either ASR or speech synthesis or both. This thesis concerns only the ASR portion (i.e., input mechanism) of the I/O subsystem and not the speech synthesis (output). Figure 3.1(b) depicts the voice recognition system we are dealing with. A thorough understanding of the human, his tasks, and the interface tasks done by the ASR component will establish the need and advisability of interjecting speech into the system. It is possible voice input is not appropriate to the application being considered. [Ref. 15: p. 37]

If a system developer wants to incorporate ASR technology, he will have several options to pursue concerning the types of ASR equipment. One can select between a discrete speech or a continuous speech recognition system. One can then choose a speaker dependent or speaker independent system.

A discrete speech, also called isolated utterances, system recognizes individual words or short phrases spoken with distinct pauses between words. This results in a stilted method of speaking that is not natural for a person. In a continuous voice recognition system, specific words or phrases are recognized by the system in the context of normal speech. The speaker talks as if in a conversation and without pauses to isolate the desired input. This mode of ASR is far more natural to the speaker than is the isolated utterances mode required by a discrete system.

The next choice for the developer is speaker dependence or independence. This choice is based on factors such as the number of different speakers to use the speech system, training time available to new speakers, and the proposed use of the voice recognition system. For a speaker dependent system, the individual user must train the system with a finite vocabulary of words so it will recognize the speech input of that user. The size of the vocabulary depends on the manufacturer. Once a user has trained the system, his voice patterns are saved and will be available to the user at any time. It is not necessary to retrain the system each time the person accesses the system. The user creates the vocabulary before training begins. Since the system is dependent upon the speaker, each vocabulary can be changed to meet the needs of the individual user.

A speaker independent system does not need to be trained by users. Anyone can use it at anytime. The limitation is that the vocabulary is set and cannot be modified to suit personal needs. Currently, speaker independent, continuous speech systems are not yet readily available for integration into existing applications.

B. SPEECH RECOGNITION EQUIPMENT AT NPS

For this thesis we considered several types of ASR systems available at the Naval Postgraduate School (NPS). One system is the Threshold Technology, Inc. Model T600 voice recognition system. With memory enhancement, the system can handle up to 256 discrete utterances. An utterance is a word or phrase lasting from 0.1 to 2.0 seconds. [Ref. 16: p. 28]

There are also two continuous speech systems, the Verbex 3000 and Verbex 4000, developed by the Verbex Company. The 3000 can handle up to 360 utterances divided into as many as 20 separate vocabularies. When a person uses one vocabulary, input can be continuous. If more than one vocabulary is involved then the user must pause as the system switches to the appropriate vocabulary. Each vocabulary is a finite number of words. The user must program his own applications, thereby making him independent of the customer engineering services which Verbex provided. The Verbex 3000 unit is a large, bulky system (9 to 10 cubic feet) so we did not consider this a viable option for testing with JAMPS II.V. [Ref. 16: pp. 29-30]

The Verbex 4000 is a smaller, more portable system. It is used in conjunction with an IBM Personal Computer. The vocabulary is limited to approximately 100 utterances. The user builds each vocabulary by using the computer's line editor. The vocabulary is then translated by the Verbex-supplied software for subsequent use by the 4000. Although the vocabulary capacity is 100 utterances, a complex vocabulary can significantly reduce the system's word capacity.

A third continuous speech system recently became available at the NPS. We chose this ASR system, the Texas Instruments (TI) TI-SPEECH, for our thesis work. It can be installed in the TI Portable Professional Computer or an IBM Personal Computer, both of which are available at the NPS. The next chapter will detail the system, its use, and the development of vocabularies. There were four basic reasons why we chose the TI-SPEECH:

- IBM compatibility;
- 450-word grammar in 9 separate vocabularies;
- ease of training;
- portability.

C. RELATED STUDIES

There have been a great number of studies conducted at the NPS in the area of voice recognition. The reports include the performance of naive versus practiced speakers, the effect of feedback on users of ASR equipment, and the effects of stress and changing environments on the user of various recognition systems. A more complete list of the studies can be found in Appendix B.

One particular study explored the use of continuous speech, discrete speech, and keyboard input to an interactive naval warfare simulation. The authors of the study and a follow-on report did not consider the results conclusive.

However, the results did demonstrate a definite advantage of continuous speech input over the discrete mode [Ref. 16: p. 2]. *This, combined with the ASR capabilities of the speech-equipped TI computer discussed in the next chapter*, impacted on our decision to select a continuous speech system, particularly the TI system.

D. ADVANTAGES AND DISADVANTAGES OF SPEECH I/O

As one can see in Tables 2, 3, and 4, there are both pros and cons to using ASR as an input device [Ref. 15: p. 36]. These tables also list the advantages and disadvantages of speech synthesis as an output technique. Since we are concerned only with speech input, ASR items of interest are highlighted below. This is not to say speech synthesis is not an extremely important aspect of speech technology. Instead, it is beyond the intended scope of this thesis.

1. Engineering Advantages/Disadvantages

The first two advantages, speed and accuracy, are keys to the application of ASR to any system. They are especially important in a C2 subsystem such as JAMPS *II.V*. Speed and accuracy emphasize our goal of reducing the time to pass on decisions and shortening the timeline within the C2 process. The additional advantages listed can enhance an existing C2 system by integrating ASR technology.

The disadvantages can be dealt with by controlling the environment. People can reduce ambient noise and adjust the physical conditions under which the system will operate. The third disadvantage does not apply, since JAMPS is a message preparation system and record copies of the input are made automatically at the operational facility. Finally, the use of a microphone is not necessarily a disadvantage because it focuses the attention of the user on the video terminal and reduces the amount of external noise that will intrude on the ASR system. When an operator wears a head-mounted microphone, he frees his hands for other tasks.

TABLE 2

ENGINEERING ADVANTAGES AND DISADVANTAGES OF ASR INPUT

ADVANTAGES
Can be faster than other communications modes.
Can be more accurate than other communications modes.
Compatible with communications systems (telephone).
Can reduce manpower requirements.
DISADVANTAGES
Possible interference from noise, distortions, and competing talkers.
Physical conditions (vibrations and physical orientation of speaker) may change speech patterns.
No permanent record of speech (unless explicitly recorded).
Microphones needed for speech input, and acoustic speakers needed for speech output.

2. Pyschological Advantages and Disadvantages

When one deals with the interface between man and machine, the psychological problems take on an even greater significance than one might realize. The closer to natural speech the ASR system conforms, the easier it will be to implement the speech recognition techniques. Speech is considered the most natural form of communication. Being able to "talk to one's computer" could ease some of the reluctance an inexperienced computer user might encounter.

The psychological disadvantages parallel the advantages. The same computer fear that might be eased by the

TABLE 3

PSYCHOLOGICAL ADVANTAGES AND DISADVANTAGES OF ASR INPUT

ADVANTAGES
Most natural form of human communication.
Best for group problem solving.
Universal among humans.
Can reduce visual information overload.
Increases in value when person is engaged in complex thought processes.
DISADVANTAGES
Speech is not private; others may eavesdrop.
Psychological changes (stress) may change one's speech characteristics.
Speech synthesis may interfere with other aural indicators.

ability to speak commands to a machine could be intensified instead. Humans tend to feel awkward when speaking to an inanimate object. However, successful operation of the system in the initial stages should ease the problem. As ASR becomes more prevalent, it may well become the normal mode of input. The environment in which the JAMPS II.V will be used (such as tactical C2 centers) will negate any problems encountered as a result of the lack of privacy associated with speech. These facilities and the associated messages are not intended to be private to those people who work there.

TABLE 4

PHYSIOLOGICAL ADVANTAGES AND DISADVANTAGES OF ASR INPUT

ADVANTAGES
Requires less effort and motor activity than other communications modes.
Frees the hands and eyes.
Permits multimodal operation.
Feasible in a darkened area.
Is omnidirectional; doesn't need direct line of sight between user and ASR system.
Permits operator mobility.
Contains information on identity and emotional state of the speaker.
Contains information on physical state of the speaker.
Simultaneous interaction with man and machines.
DISADVANTAGES
Prolonged speaking may cause fatigue, which may in turn change speech characteristics.
Illness may change speech characteristics.

3. Physiological Advantages and Disadvantages

The physiological advantages work toward easing the stress and fatigue on an operator making him more versatile in a C2 environment. As for the disadvantages, any factor which reduces these negative factors will aid and improve the C2 system. Both ASR technology and the operators themselves should recognize and compensate for these adverse factors.

E. SUMMARY OF ASR

In this chapter we talked in general terms about voice recognition and what ASR systems are available at NPS. We also wanted to address some of the pros and cons of using voice recognition in a C2 environment. These are valid points to weigh when considering the interface of ASR technology with a C2 system like JAMPS II.V. The next chapter will deal with the specific operation of the TI-SPEECH on the TI Portable Professional Computer.

IV. TI-SPEECH

A. INTRODUCTION

As mentioned in the previous chapter, there were several voice recognition systems available at the Naval Postgraduate School for our use with this thesis. The system we selected was the Texas Instruments Portable Professional Computer (TIPPC) with the Texas Instruments' SPEECH (TI-SPEECH) system installed.

Earlier studies have demonstrated a significant advantage of continuous speech recognition over discrete speech recognition for data input. The TI-SPEECH system can be operated in either a continuous recognition mode or a discrete recognition mode.

The TI-SPEECH has a vocabulary that far exceeds any other system that was available. Each of the 9 vocabularies has the capacity for as many as 50 utterances.[3] However, each vocabulary can only be 32 seconds in length. If the user develops his vocabulary with very short utterances, then the 50-utterance limit will be achieved. If the utterances are longer, then the time limitation may be exceeded before the utterance limitation. The system was designed to enable the user to link up to 9 vocabularies together, thus a total of 450 utterances are available to the user.

The ease with which the ASR system can be used was also a factor in our selection. Texas Instruments' use of menus and pop-up windows has simplified the procedure to the point that entry of a new vocabulary takes much less time than the other ASR equipment at this school. A user with very little experience can enter, train, and refine a 50-word vocabulary in less than an hour.

---

[3]We define an utterance as the word or phrase the user expects the ASR system to recognize and respond to.

Since the JAMPS is meant to operate on mobile microcomputers, the portable TI system's (with TI-SPEECH imbedded) physical characteristics and computing capability make it a viable option to consider in order to run the JAMPS software with voice recognition.[4] With the exception of the Verbex 4000, the other ASR systems available at the Naval Postgraduate School are large and awkward. The degradation in mobility would offset any gains provided by using speech input. The Verbex 4000, although smaller and lighter than the Verbex 3000, could integrate well with JAMPS. However, the limitations on the vocabulary size and the difficulty in creating vocabularies offset the size advantage.

The following sections will describe the hardware of the TIPPC, the software involved in the speech recognition system, and some guidelines for developing, entering, and training a vocabulary. In Appendix C is a user's guide that provides a basic step-by-step procedure for using the speech recognition system and bringing up JAMPS.

B. TEXAS INSTRUMENTS PORTABLE PROFESSIONAL COMPUTER

The basic TIPPC consists of four standard parts:

- keyboard;
- operating instructions;
- diagnostic software;
- system unit.

Texas Instruments provides these basic components for each computer [Ref. 17: p. 1-3]. There are many other options that can be added to the TIPPC; some are hardware related

[4] Remember, our primary purpose is to investigate the feasibility of using voice recognition as an input medium to JAMPS. Using TI-SPEECH meant that the JAMPS software had to be loaded onto the TIPPC, which is what TAFIG and MITRE helped us do. At this time, we feel that the TIPPC with TI-SPEECH installed, is an option worth pursuing for the running of JAMPS. We stress that more research is necessary in order to substantiate or refute our thesis. Our last chapter, CONCLUSIONS AND RECOMMENDATIONS, lists some topics to consider for future research before any decision would be made concerning the use of voice recognition with JAMPS.

and some are software related. The voice system (TI-SPEECH) is one such option. The voice system incorporates both hardware and software and will be detailed later in the chapter.

1. Keyboard

The keyboard is similar to the one found on an electric typewriter. The keyboard, as you see in Figure 4.1, is divided into four major areas: [Ref. 18: p. 1-1]

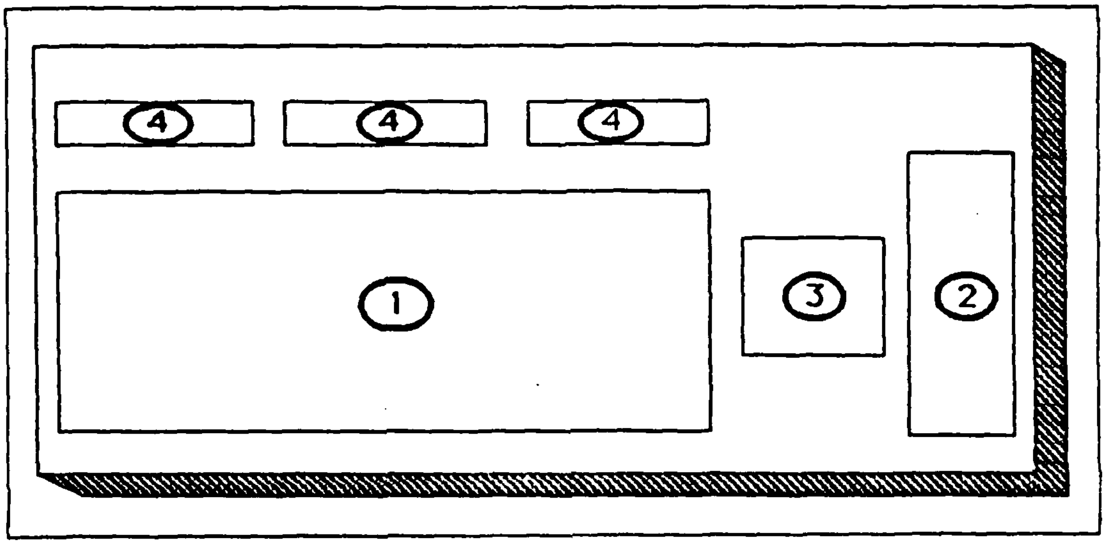


Figure 4.1 TIPPC Keyboard Layout

- main keyboard area - similar to a typewriter and used to enter alphanumeric data;
- numeric keypad - used as a calculator and for numeric entry;
- five key cluster - used to control display cursor movement;
- function keys - 12 keys arranged in 3 groups of 4; these keys are programmable, and their functions can be changed from application to application.

The keyboard also contains a microprocessor which converts keystrokes into character information and conducts keyboard diagnostics each time the system is powered up [Ref. 18: p. 1-1]. There are other special purpose keys present on the keyboard; some will be used and some will not. As it becomes necessary to use a special key, we will define its purpose and indicate its position on the keyboard.

2. Operating Instructions

The operating instructions provide information on the initial set up and operation of the TIPPC. Included are chapters on problem solving, hardware options available, shipment or movement of the computer, and several appendices. The operating instructions are invaluable to both a first-time computer user and an experienced user.

3. Diagnostic Software

The diagnostic software provided by Texas Instruments checks out the computer system. The tests run by the diagnostics are more extensive than the power-up checks run each time the system is turned on. These tests will aid in the identification of problems should the system not operate properly.

4. System Unit

The system unit consists of the following basic components:

- central processing unit (CPU);
- parallel printer port;
- power supply;
- random access memory (RAM);
- system unit board;
- 5 1/4 inch diskette drive;
- display unit.

These components make up the computer itself. In the portable model we operated, the components are all in a compact unit 19 x 17 x 8.5 inches. [Ref. 18: p. 1-2]

The CPU of the system unit consists of an Intel 8088 16-bit microprocessor, clock circuits, bus buffers and latches, bus controller, and the reset circuit. [Ref. 18: p. 2-11]

The parallel printer port contains a 25-pin female D-Type connector. Printers with centronics-compatible interfaces use this port. [Ref. 18: p. 2-16]

The power supply is a switch type, 110 watt unit with 3 output levels. It is designed to support a system equipped with every hardware option available for installation into the computer. [Ref. 18: p. 1-2]

The basic computer system contains 64 kilobytes of dynamic RAM. The computer we used already had its RAM expanded to the maximum of 768 kilobytes by adding optional expansion memory boards. [Ref. 18: p. 1-3]

The system unit board, also called the motherboard, is the heart of the computer system. It is divided into two sections. First, the systems support section contains some key components such as the keyboard port, CPU, input/output system, interrupt and memory systems. As we mentioned earlier, the CPU consists of an Intel 8088 16-bit microprocessor. Our system has been enhanced with the optional Intel 8087 numeric data processor (also known as a numeric co-processor) [Ref. 18: pp. 2-10 to 2-11]. The second section of the motherboard is for the expansion bus. This section provides the space for installing the various option boards available for the TIPPC [Ref. 18: p. 2-56].

One 5 1/4 inch diskette drive is standard with the TIPPC. The drive is a mass storage device for reading from or writing to a standard 5 1/4 inch removable diskette. The diskette can store approximately 320 kilobytes of data.

There are many hardware options available for installation onto the TIPPC. The following options were added to the computer system we used: [Ref. 18: p. 1-3]

- expansion memory boards - total RAM is 768 kilobytes;
- numeric co-processor;
- Winchester (hard disk) drive - the Winchester disk drive and controller are available with 5 or 10 megabyte capacities. The system available to us contained the 10 megabyte drive. It is installed internally to the computer system in the space set aside for a second 5 1/4 inch diskette drive.

- synchronous-asynchronous communications board - this board permits communications through a standard RS-232-C interface. Asynchronous data rates from 50 bits per second (bps) to 19,200 bps are supported by the communications board.

There are other options available but are not installed. These include internal modems and a graphics control board. They are not applicable to our thesis and will not be discussed.

C. AUTOMATIC SPEECH RECOGNITION SYSTEM

The entire TI-SPEECH system encompasses many aspects of voice processing. These include voice recognition record and playback, text- to-speech and telephone management capabilities. The integration of all these capabilities raises interesting possibilities for use with JAMPS, but that goes beyond the scope of this thesis. We will concentrate on the voice recognition aspect, Vocabulary Manager (VM), of the TI-SPEECH system. The VM, which we describe later, is coupled with a Transparent Keyboard(tm) feature that overlays other applications.[5] When the Transparent Keyboard recognizes a voice command, it passes the appropriate command along as if the keyboard had been used.

1. Hardware

The TI-SPEECH hardware consists of several items. These include the following: [Ref. 19: p. 1-2]

- microphone and speaker;
- TI-SPEECH Diagnostics, Algorithms and DSR Diskette;
- TI-SPEECH board;
- TI-SPEECH Installation Guide.

The microphone and speaker are external items used as input/output devices. They can be replaced by an optional headset. The headset includes a microphone for input and a single earpiece for output. The headset frees the user's hands for other operations. If voice synthesis

[5]Transparent Keyboard(tm) is a trademark of Texas Instruments, Inc.

is used, the earpiece provides privacy. We used the headset for our test runs.

The installation guide is exactly what the name implies. It provides the detailed instructions needed to install the TI-SPEECH board into the computer. It also provides a description of the diagnostic routines available with the hardware kit.

The TI-SPEECH Diagnostics, Algorithms and DSR Diskette might be considered software, but since Texas Instruments considers it part of the hardware kit, we will discuss it here. The diagnostic routines test the speech board and optional telephone interface board, if installed. If more than one speech board is installed, each one is tested in sequence. The routines check the overall operation of the speech board, and a final test checks different combinations of input and output devices. [Ref. 19: p. 5-7]

The speech algorithms implement the capabilities of the speech board. The software is executed by the TMS320 (see below). There are eight different algorithms that deal with the speech processing available. They include algorithms dealing with text-to-speech, analysis/synthesis, recognition/synthesis, dial telephone, and continuous recognition [Ref. 19: p. 1-8]. We will be using only the continuous recognition algorithm. This algorithm provides for the recognition of a user's isolated (discrete), connected, and continuous spoken phrases. These three modes provide the user with more flexibility in customizing his speech input to specific jobs or applications [Ref. 19: p. 1-12].

The device service routine (DSR) enables the disk operating system (DOS) software to access the capabilities of the TI-SPEECH hardware. The DSR provides the software interface between the speech hardware and the application. The DSR must be installed into computer memory each time a user wishes to use the speech hardware. Once installed, the

DSR remains available unless the system is rebooted, in which case it must be reinstalled. [Ref. 19: p. 1-6]

The TI-SPEECH board is the key to the speech processing system. See Figure 4.2 for a general overview of the speech board architecture. [Ref. 19: p. 1-4]

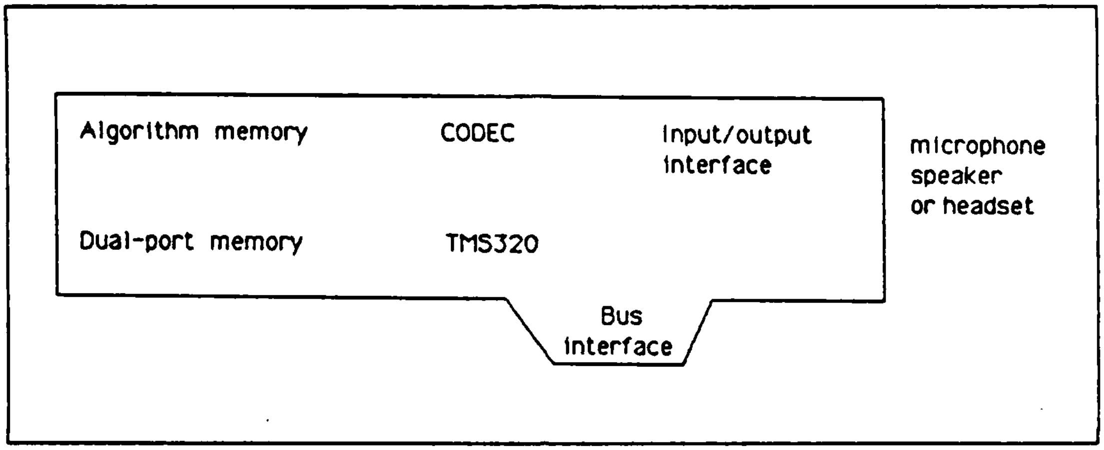


Figure 4.2 TI-SPEECH Board Architecture

The TMS320 is a 32-bit microprocessor that executes 5 million instructions per second. It is a special-purpose microprocessor for digital signal processing. The speech algorithms are executed by the TMS320. The algorithm memory uses eight kilobytes of the TI-SPEECH board RAM. The current algorithm in use is stored and available to the TMS320 for execution. The dual port memory stores speech data during speech processing. It consists of 32 kilobytes of speechboard RAM. Here is where the specific speech templates are held when using speech recognition. These templates are the voiceprints created by each user for his specific vocabularies. The memory is referred to as dual port since it can be addressed by the bus interface of the computer as well as the TMS320. Dual-porting enables the TMS320 to transfer speech data more quickly. [Ref. 19: p. 1-4]

The input and output interface is the circuitry that connects the speech board with the input/output devices, headset and speaker (or microphone). It enables the user to select specific devices and to set output volume and input gain. [Ref. 19: p. 1-5]

The CODEC is a standard analog-to-digital (AD) and digital-to-analog (DA) converter. It changes the analog inputs to digital signals for processing by the TMS320. It then translates the microprocessor's digital response to analog signals that are output to the user. [Ref. 19: p. 1-5]

2. Software

There are several types of software available when using the TI-SPEECH system:

- NaturalLink(tm);
- MS(tm)-DOS operating system;
- TI-SPEECH Vocabulary Manager.[6]

The NaturalLink software allows the novice user to more easily access MS-DOS. It is not necessary for the execution of the speech recognition software. Appendix C, the user's guide, will show the user how to bypass the NaturalLink and execute the Vocabulary Manager programs without it. One can find further details on the use of NaturalLink in the Texas Instruments' manual, "NaturalLink Access to MS-DOS," TI Part Number 2243754-0001.

MS-DOS is the disk operating system (DOS) provided with the TIPPC. The VM software is designed to use MS-DOS or the IBM Personal Computer DOS. Our TIPPC is using MS-DOS Version 2.13. The disk operating system is a collection of basic computer programs designed to simplify tasks you need to use with your computer from time to time. Basically, it acts as the manager/controller for your computer.

---

[6]NaturalLink(tm) is a trademark of Texas Instruments, Inc. MS(tm) is a trademark of Microsoft Corporation.

The Vocabulary Manager (VM) is the software portion of the TI-SPEECH system that enables a user to give voice commands to his computer. Using VM, an operator can create and store his own vocabularies for specific applications, refine and test created vocabularies, link several vocabularies together or use and modify the sample application vocabularies provided with the TI-SPEECH system. [Ref. 20: pp. 1-2 to 1-3]

The VM software kit contains a user's manual, the TI-SPEECH VM guide with three ring binder and tabs, two Vocabulary Manager diskettes, and four Vocabulary Manager tutorial diskettes. The software kit, when combined with a computer and the hardware kit previously discussed, make up the Automatic Speech Recognition system.

The tutorial diskettes are excellent tools for familiarizing a user with the basics of using the VM. It is recommended the first-time user utilize the tutorial to get started with the ASR system [Ref. 20: p. 1-2]. If the tutorial is not available, our user's guide found in Appendix C will provide the basic commands to enter, train, and begin using a vocabulary.

When the user first brings up the Vocabulary Manager, the main VM menu will appear. Figure 4.3 shows a general outline of the main menu.

Section 1 is the menu area in which the commands the user selects are built. When a command is complete, it will prompt the user to press the ENTER key found on the lower right hand corner of the numeric keypad described in the hardware section of this thesis.

Section 2 is the area in which you can find the command segments. A person selects a segment by using the control arrows to move the highlighted area to the desired segment and pressing the RETURN key. Section 3 is a listing of available vocabularies. The arrows, section 4, indicate

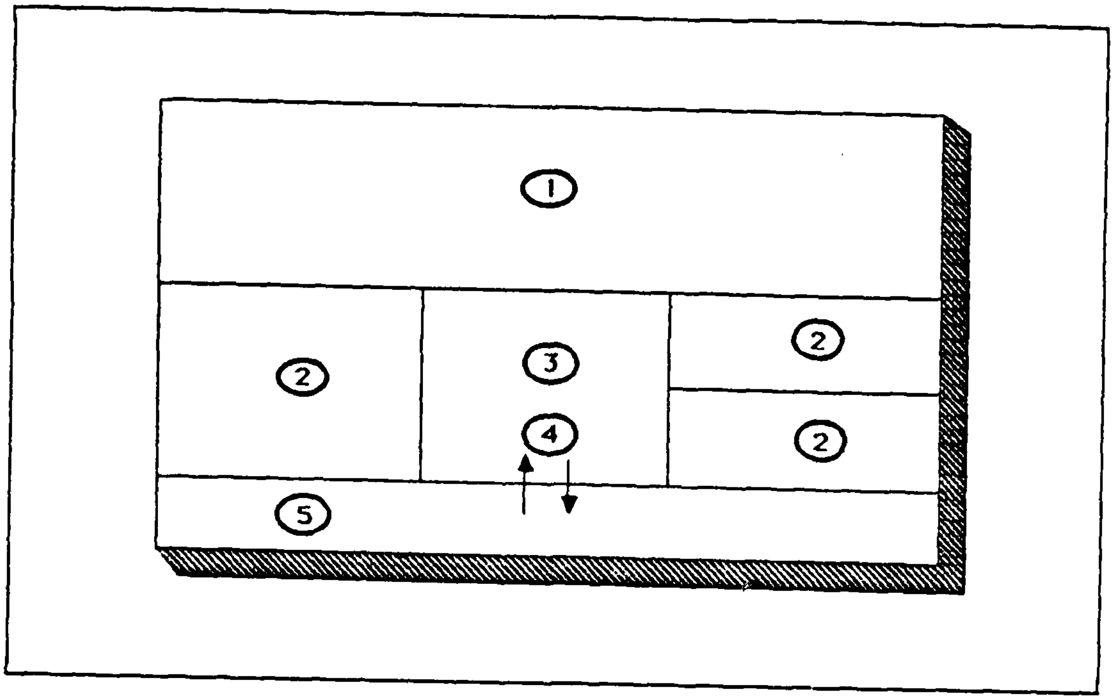


Figure 4.3 Vocabulary Manager Menu

there are more items above or below the window which the user can bring into view.

Section 5 shows the function keys available at any particular time during the use of the Vocabulary Manager. Occasionally pop-up windows will appear over the menu. These windows display help messages or prompt the user to input more information [Ref. 20: p. 2-4].

The cursor control keys are between the numeric keypad and main keyboard area (See Figure 4.1). Use the cursor arrows to move the highlighted area from command segment to command segment. The RETURN key is found on the right side of the main keyboard and is used to select a highlighted segment. The ENTER key, as mentioned above, will execute a command once it is built.

D. BUILDING A VOCABULARY

The first thing a user must do is decide to which application he is going to apply voice recognition. In the case of this thesis we selected the JAMPS software. Once the user is familiar with the commands needed to operate the software, he is ready to begin building his vocabulary.

The next step is the selection of phrases or utterances. These are the words the user will say when executing his software. The phrases or words can be anything the operator chooses, but some things should be kept in mind. The words should relate in some way to the commands they will be associated with; this will make them easier to remember during use. Multi-syllable words are better since they give the computer more sounds to recognize. Homonyms, such as right and write, should be avoided for obvious reasons. Words that appear within other words, like the six in sixteen or sixty, should also be avoided. Consideration of these points will make it easier to design a vocabulary and have it recognized with a minimum of errors. [Ref. 20: p. 3-3]

Once the person forms the phrases, then he must match the keystrokes (maximum of 40 characters) needed for the execution of the commands associated with phrases.[7] When the complete list of phrases and keystrokes is done, the user is ready to enter them into the computer. Appendix C provides the commands and steps for the entry.

After the user has entered the vocabulary, he must create his voiceprints by "training" the speech system to recognize his voice. To enhance the effectiveness of the training session and subsequent recognition in the operating environment, several guidelines should be considered. You should use the same input device each time for the same

---

[7]Note that "phrases" is the left hand column displayed by the Voice Manager. "Keystrokes" is the right hand column. Phrases are your spoken words or commands, while keystrokes are the corresponding computer response.

vocabularies. If training is done using a headset then the headset should be used during operations. The input device should be placed in the same location each time it is used. Insure the input device is on before beginning the training session. Turning it on and off during a session could create problems. Training should be done in an environment that closely approximates the operating environment. If the operating environment is very noisy, the voice patterns should be created in a quiet environment and subsequent training done in a noisy environment approximating the operating environment. Also, the user should use a normal tone of voice. The same speech style will help the user reach consistent recognition. [Ref. 20: p. 3-6]

As mentioned earlier, the system has the capability of recognizing continuous speech. To enable this mode the user must first create the voiceprints using only the phrase. Then the voiceprints should be refined using the phrase to be recognized in a sentence. Sentence content is not important. The idea is to get the computer accustomed to recognizing the phrase in sentences with other words. [Ref. 20: p. 3-6]

The ASR system user can edit a previously created vocabulary with a series of commands. The commands become available to the user when he selects the command "set up and train a vocabulary" from the main menu. The options that appear permit the user to further customize and refine his vocabulary. [Ref. 20: pp. 3-10 to 3-11]

The individual must test his voiceprints when he has finished creating them. As the voiceprints are tested, the VM provides feedback which is of great value to the user in enhancing error-free recognition. As the user speaks each phrase, a pop-up window, as in Figure 4.4, will appear.

If TI-SPEECH does not recognize the phrase, the message in the window will indicate such. The user should repeat the

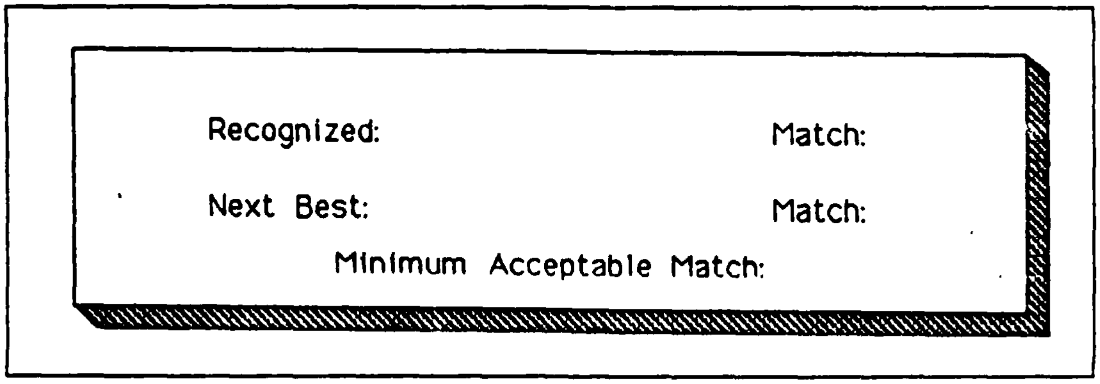


Figure 4.4 Pop-up Window for Vocabulary Testing

phrase; and if recognition still does not occur, the user should recreate the voiceprint. If the phrase is recognized, the pop-up window will show a match number ranging from one to ten. Seven to ten indicates a good match; four to six is a fair match; and one to three is a poor match. A poor match means recognition will not be consistent and the voiceprint should be further refined to insure consistent recognition. A fair match means recognition will occur, but errors will also be present; further refinement might be necessary.

The next best match gives the user an indication that another utterance has a voiceprint similar to the actual spoken phrase. If the match number in the next best match differs by only a value of one or two, then the user should change one of the phrases to avoid possible recognition errors. Usually the voiceprints will be sufficiently unique so that no next best match will be given.

Finally, the minimum acceptable match shows the level at which the VM will provide recognition. If the person sets the level to one, the speech system will recognize all matches equal to one or greater. If the individual sets the level to four, TI-SPEECH will recognize matches greater than or equal to four. You can adjust the minimum acceptable match with the "optimize recognition" command. [Ref. 20: p. 4-4]

Each vocabulary a user enters can hold a maximum of 50 words. A more definitive limitation is the fact each vocabulary is limited to 32 seconds of phrase time. If short phrases are used, it is possible to reach the 50 phrase maximum. However, if a person incorporates longer phrases, then the time allocated will be expended before reaching the 50 phrase limit. In the case of some complicated applications, the user must exceed the 50 phrase limit to create all the commands he needs. This is not a problem since VM provides the capability to link as many as nine vocabularies together.

There are some points to keep in mind when linking together a group of vocabularies. First is the amount of RAM available. Each vocabulary occupies 26 kilobytes of RAM when it is full [Ref. 20: p. 4-6]. Since our computer system had been fully enhanced to 768 kilobytes of RAM and we linked only three vocabularies, we did not encounter a memory problem.

If you are going to use more than one vocabulary, you must initially load the ones you need; otherwise the computer must be restarted to make additional vocabularies available. The user should also keep in mind the VM will only recognize phrases in the "active" vocabulary. This means that the same phrase may be placed in different vocabularies, but with different keystrokes associated with it. The user must remember which vocabulary is active, since the TIPPC provides no indication. [Ref. 20: p. 4-6]

When a user has more than one vocabulary available in memory, he has one of two ways to move from one vocabulary to another. First, he may use the main menu. This necessitates leaving the application in which he is working, changing the active vocabulary, and then returning to the application. This is not a difficult process but it does take time. [Ref. 20: p. 4-7]

The other method is through the use of key switch words which enable the active vocabulary to be changed by a voice command. Insertion of the switches adds a little time to the preparation of a vocabulary [Ref. 20: p. 4-8]. However, the user can change vocabularies without leaving his application program. In the case of JAMPS this can be a significant time saver.

E. SUMMARY

In this chapter we introduced you to the Texas Instruments Portable Professional Computer and TI-SPEECH. This was a necessary overview of the hardware and software associated with the IBM-compatible microcomputer that the JAMPS software can run on, and the voice recognition system which can execute the JAMPS software. We also included some helpful information about creating, training, and using vocabularies. More details can be found in Appendix C. And if the reader wants even more information, the references for the TI manuals are in the List of References.

V. A MEASURE OF PERFORMANCE FOR MESSAGE PREPARATION

This section covers the development of a measure of performance (MOP) which can be used in preparing messages. The goal is to quantitatively compare the automated method of message preparation (JAMPS) versus the manual method of preparing JINTACCS messages. Specifically, this MOP focuses on the time element involved in preparing messages up until time of transmittal. Time is a crucial factor in the C2 process when some military force must act before a given event takes place. JAMPS is designed to save time over manual message preparation so a quantitative comparison is essential in determining JAMPS usefulness. Since this performance measure is simple and generic, it could also compare JAMPS using voice recognition as an input medium for some portions of hard-copy messages to JAMPS using just the keyboard for input.

A. WHY AN MOP?

In doing this research, we had a choice of developing a measure of performance or a measure of effectiveness (MOE) to characterize the system involved in the message preparation process. As stated in the 1985 Command and Control Evaluation Workshop, performance and effectiveness measures should have the criteria [Ref. 21: p. 6-13] listed in Table 5. The only criterion that was suspect in choosing a measure was independence. One could conceivably have a measure which affects another measure. For example, let's say one measure for a communications system is the average number of formatting errors per message prepared. Let's further assume a message cannot be transmitted with formatting errors, so an operator must correct these errors. What if another measure of this system is the time needed to prepare a message for transmission. In this case the format errors affect the time for preparation as a measure. Hence these measures could not be mutually exclusive.

TABLE 5

DESIRED CRITERIA FOR PERFORMANCE AND EFFECTIVENESS MEASURES

CHARACTERISTICS	DEFINITION
Mission Oriented	Relate to force/system mission.
Discriminatory	Identify real differences between alternatives.
Measurable	Able to be computed, or estimated.
Quantitative	Able to be assigned numbers, or ranked.
Realistic	Relate realistically to the C2 and associated uncertainties.
Objective	Defined independent of subjective opinion. (Some measures cannot be objectively defined).
Appropriate	Relate to acceptable standards and analysis objectives.
Sensitive	Reflect changes in system variables.
Inclusive	Reflect those standards needed by the analysis objectives.
Independent	Mutually exclusive with respect to other measures.
Simple	Easily understood by user.

Using the 1985 Command and Control Evaluation Workshop as a guide, we chose to use an MOP instead of an MOE in this analysis. We are looking at a certain portion (i.e., subsystem) of a larger C2 system. Specifically, our subsystem is the equipment and procedures used for message preparation which occurs in an operational facility. Our larger C2 system is the entire tactical command and control system and its accompanying C2 process associated with a specific mission, such as air defense or fire support. By choosing an MOP vice MOE, we can clearly bound the C2 system we are analyzing. We set these boundaries based on the application of our analysis, and the process we analyzed takes place within the limits of the C2 system. We have delineated the system being studied from its environment [Ref. 21: p. 8-4]. In this way we can more easily model the message preparation process. The next section outlines the system of interest and its boundaries.

B. MODEL OF THE MESSAGE PREPARATION PROCESS

In drawing a model of the message preparation process, we had to concisely depict the steps involved in preparing a message for transmission. Figure 5.1(a) shows this current process. Figure 5.1(b) shows the JAMPS process.

1. Manual Preparation

In a tactical C2 facility, the message process starts with some stimulus to send a message. (See Figure 5.1(a) for a model of the message preparation process). A stimulus may result from a change in the battlefield environment or a need to redirect forces. Because of this stimulus, a commander (more likely his staff) will hand-draft a message on a joint message form, which must then be typed accurately to meet the stringent JINTACCS formatting. The drafter will normally review the typed message, check for errors, and either return the message for corrections or give the message to the releasing authority. The releaser usually reviews the message, checks for mistakes, and then signs the message for release if there are no changes. Once the message is ready to be transmitted, a runner must carry

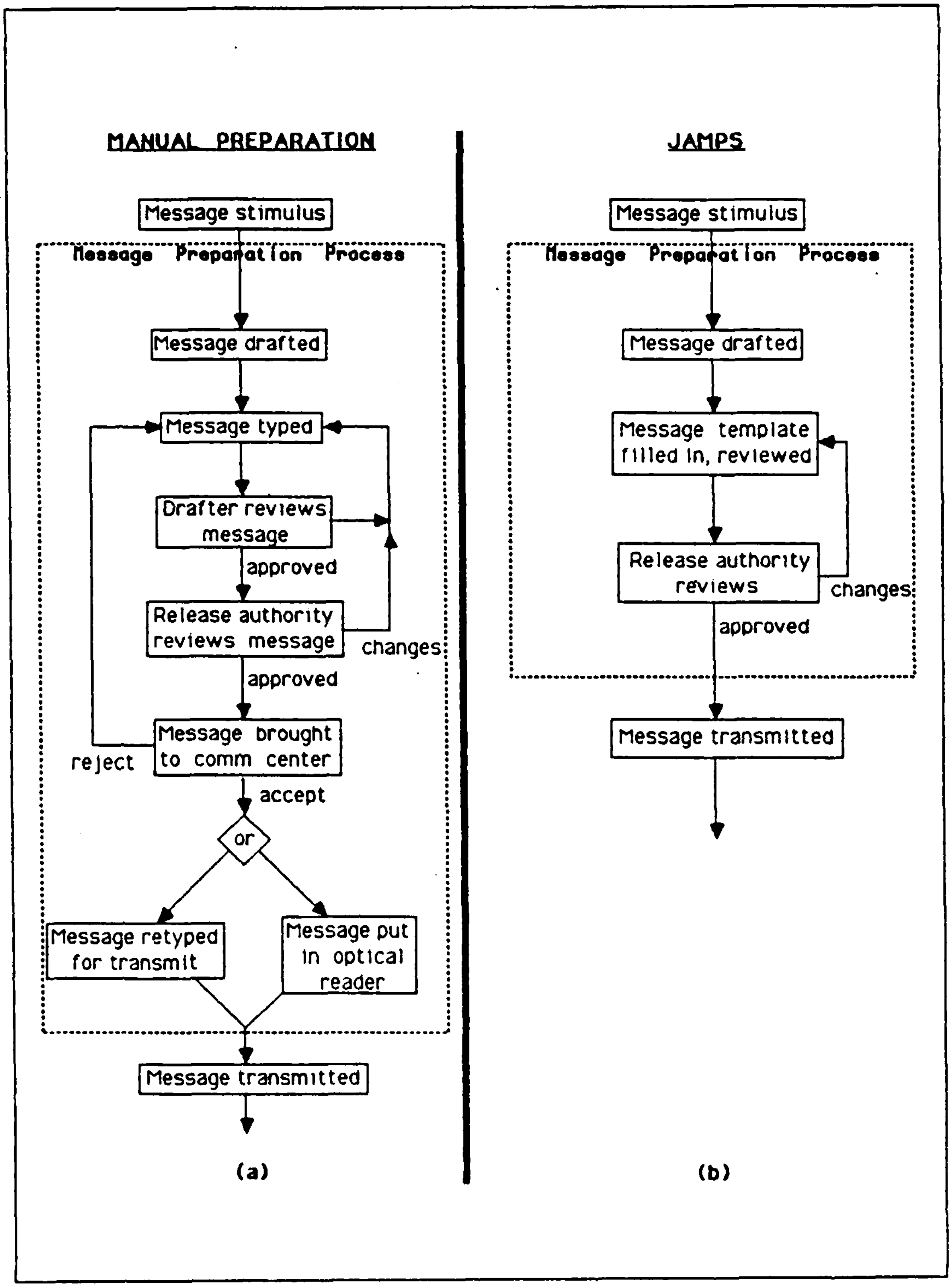


Figure 5.1 Message Preparation Model

the message to a communications center. Here at the center, one of two things will happen. The communications center's personnel will either accept the message and log (record) it in or reject it. If they accept the message, then they are responsible for retyping the message or putting it through an optical reader for subsequent transmission over external military communications. If the communications center rejects the message for errors or some other reason, then the operational facility must redo the message. More time is lost.

There are certain administrative procedures inherent in the manual preparation model that take additional time. For instance, once the message arrives at the communications center, the message must first be accepted and then logged in by the center. The operational C2 facility must also record the message information. And should the communications personnel tear the message form or make an error, then they would need to retype the joint message form. We did not reflect these added steps as separate blocks in the model, but one should be aware of them.

This model is generally the process followed by military units when in joint tactical operations. The specific procedures, such as drafting or reviewing, may vary somewhat amongst operational facilities.

2. JAMPS Preparation

Looking now at Figure 5.1(b), one sees that the JAMPS message preparation model starts with the same stimulus for a message. Because of this stimulus, someone within the C2 facility hand-drafts a joint message form. This is where the steps vary from the manual preparation model. A JAMPS operator will use the draft message to complete a fill-in-the-blanks message template on his video terminal. The drafter can review the message on the screen or the operator will give the drafter a print-out copy of the message ready for transmittal. If any changes are necessary, the JAMPS operator can quickly recall the message from the stored messages. The next step is for the releasing authority to review and approve the message. When approved, the message is transmitted directly by the JAMPS workstation serving as the communications gateway.

Comparing the two models of Figure 5.1, one sees fewer steps in the JAMPS process. A tactical commander can conceivably save valuable time in his C2 process by using JAMPS. His operators will not need to type messages from scratch. Nor will runners need to carry a message to a communications center where more time is necessary to send out the message. But the question still remains, "How much time can be saved?"

C. ASSUMPTIONS AND LIMITING FACTORS

In developing an MOP, we had to narrowly define what we were analyzing. We made certain assumptions, which can also be limiting, in order to meet these desired criteria for measures from Table 5:

- mission oriented;
- discriminatory;
- measurable;
- quantitative;
- realistic;
- objective;
- appropriate;
- inclusive;
- simple.

Our assumptions, as they relate to the message preparation models in Figure 5.1, are as follows.

(a) The two competing systems which model message preparation are treated as "black boxes" due to the boundaries.

(b) We considered all other environmental factors as being equal for the two systems. For instance, both systems are in the same tactical scenario, they both use the same external military communications and switches for transmission, the message stimuli are the same for both models, and we are dealing with similar tactical C2 facilities.

(c) The people within the systems are qualified in their jobs.

(d) The like C2 facilities are handling like messages (i.e., message types and precedence).

(e) The systems are functioning during the time period in question. The system set-up and check-out times are neglected. This is a limiting assumption should one try to develop another MOP or expand the analysis scope to create an MOE.

We tailored these assumptions for ease in comparing the two systems. One "black box" can be tested against another just by swapping them, all other factors being equal. With the appropriate model and assumptions having been stated, the next step is to derive a measure of performance.

D. DEVELOPING THE MOP

We used an analogy to the command and control timeline analysis, as seen in the <u>Command and Control Evaluation Workshop</u>, to develop the MOP [Ref. 21: p. 5-14]. Figure 5.2 shows the timeline for the process of message preparation. The application is straightforward. Some event in the combat environment serves as a stimulus for a message. This happens at time S. The tactical commander and his staff have a certain time by which they must respond to the event with a message. Otherwise the message will be sent too late to make any difference on the combat environment. This is time NL. The time the message is actually transmitted is time T. S, NL, and T are specific points in time, not elapsed times. T may be greater than, equal to, or less than NL, depending on how well the message preparation process is executed. For illustration purposes, we used T less than NL. The time between message transmittal and message stimulus is mp (mp = T - S). Residual time, r, is the difference between the no-later-than time and actual transmission time (r = NL - T). Here mp and r are elapsed times.

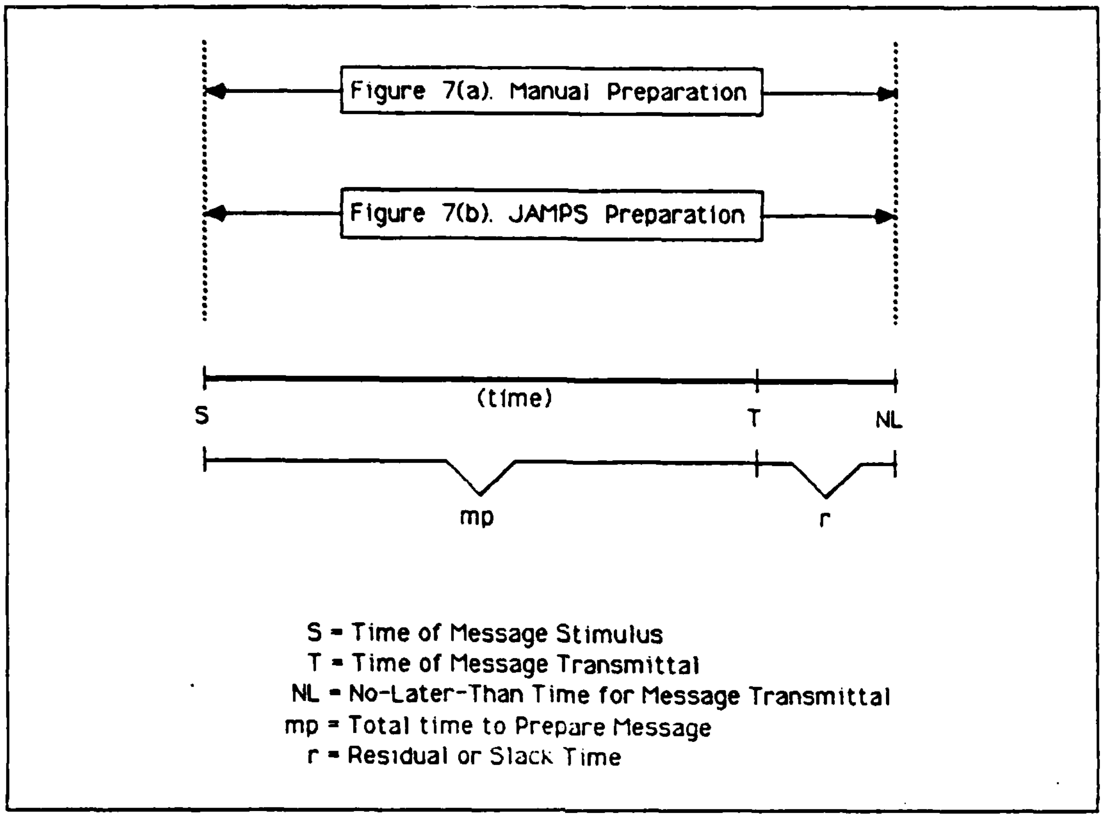


Figure 5.2 Message Preparation Timeline

Recall we are analyzing a subsystem of a larger tactical C2 system. So any process which is carried out quicker in this subsystem will mean there is more time for the larger system to go through its C2 process. The smaller the value of mp, then the more desirable the subsystem is for the message preparation process. And so, we define our measure of performance as the ratio:

$$MOP = \frac{mp}{mp + r} .$$

To compare how the JAMPS system prepares messages versus the manual method, we can use the ratio of the JAMPS MOP to the manual preparation MOP:

$$A = \frac{MOP(J)}{MOP(M)} .$$

If A is less than 1.0, then JAMPS saves time over the manual method of preparing messages. If A is greater than 1.0, then JAMPS takes more time. If A equals 1.0, then both methods of message preparation take the same amount of time. A should be less than 1.0. The inverse of A will tell one how much faster (or slower) JAMPS is in preparing JINTACCS messages then the current method. For example, if A = 0.5, then JAMPS takes one-half the time. Using the inverse of A, JAMPS is twice as fast.

E. DISCUSSION AND SUMMARY OF THE MOP

We developed this MOP using the general guidelines in Chapter 6 of the Command and Control Evaluation Workshop. One could easily expand the boundaries of the subsystem we analyzed and look at a larger system. This is equivalent to going from level 1 to level 2 in the hierarchical structure of measures. In analyzing the larger C2 system, more dimensional parameters and/or subsystem MOPs would be necessary in order to create an MOE.

This MOP is best suited for a test in a controlled environment so that one message at a time enters the message preparation process. Data collection and analysis would be simple. An exercise environment would be different since a JAMPS operator is not likely to get just one message at a time to process from beginning to end. More likely, he will get several messages, or he will get a second message (perhaps with higher precedence) while he is working on his first message. In this situation, queuing would occur and additional analysis would be needed.

We used time of message preparation as an MOP because time is so important in any command and control system. If the friendly commander can step through his C2 process faster then the enemy can execute theirs, then the friendly force has a better chance of achieving its objective. The shorter the message preparation time, the better the tactical C2 subsystem, considering just this MOP. There are many possible MOPs, such as availability, transportability, reliability, error rate, and ease of reconfiguration, which can affect a C2 system's effectiveness. The expanded analysis would be an MOE based on new C2 system boundaries, different assumptions, and different objectives.

One could analyze the C2 system in which information flows from the message stimulus, through message preparation, to the receiving operational facility's commander. The receiving entity could also be similarly modelled as a subsystem. The external communications equipment and procedures would be part of the expanded analysis since they affect information flow. Our larger C2 system would then involve the two C2 subsystems (message preparation and message reception) and their attendant processes, as well as the external communications and procedures joining the two. A possible MOE for this C2 system could be how the time of message preparation influences the receiving commander's action (i.e., percentage of his resources allocated to respond to the received message by a certain time). One could even expand this MOE to include how environmental factors, such as enemy actions and terrain, influence the C2 system.

We have demonstrated just one possible MOP based on our objective and assumptions. We have also outlined a possible MOE should one carry the analysis further. Our analytical approach was at the subsystem level, level 1, but the analysis could easily be developed at the system level, level 2.

VI. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Over the last 10 to 15 years, the U. S. military has become increasingly aware of the importance of command and control in military operations. There has been much work done concerning the theory behind the C2 process. In his C2 process model, Dr. Joel S. Lawson, Sr., has stressed the criticality of time in the SENSE-PROCESS-COMPARE-DECIDE-ACT (S-P-C-D-A) process. Relative to a given event, if friendly force A can execute its S-P-C-D-A process faster than enemy B, then A has a much better chance of defeating B. Operating at a higher C2 tempo is a crucial factor in military victory.

Along with this emphasis on C2 theory, there has been a corresponding explosive growth in the development of C2 systems. One such tactical C2 system is JAMPS. JAMPS is the software to automate JINTACCS message standards that the U. S. military will use in joint tactical operations. One of its primary design features is to reduce the time needed to prepare joint messages (hard-copy, record traffic).

Our thesis explored the initial feasibility of integrating voice recognition with JAMPS. The ultimate goal to see if voice recognition could make a JAMPS operator's job easier (and therefore reduce work stress) and save even more time in preparing messages is left for further research. As mentioned in the introduction, we had planned to use the Texas Instruments Portable Professional Computer, with TI-SPEECH installed, as the input device to the current prototype JAMPS workstation. When it became possible to implement JAMPS directly on the TIPPC, we tabled the problems generated with the interface attempt. However, we had proceeded far enough to convince ourselves it was possible.

Further work could be done to completely interface the TIPPC with the JAMPS II.V workstation. But our opinion is that it is not worth the time or effort to pursue this since these prototypes will not be fielded.

We used the TIPPC, with TI-SPEECH imbedded, as our automatic speech recognition system. JAMPS had to be loaded onto the TIPPC's hard disk before we could use voice recognition to run the software. This enabled us to prove that voice recognition can be integrated with JAMPS. Based on our research, we strongly believe ASR can make a JAMPS operator's job easier and save time in preparing messages. Our measure of performance could be used in exercises or controlled scenarios to test our hypothesis. Still more research is necessary to explore in greater detail the pros and cons of ASR with JAMPS.

B. RECOMMENDED AREAS OF FURTHER RESEARCH

Our thesis work was a preliminary investigation of using ASR with JAMPS. We offer five categories to consider in any possible future research. These are areas we think people should examine before any final judgement is made concerning the use of ASR with JAMPS. This is not an exhaustive list of possible research topics.

1. Refining the Grammar and Operating Procedures

The grammar we created was a test grammar. We were not concerned with optimizing the vocabularies for ease of use with JAMPS. A closer study can reveal the appropriate phrases and words to incorporate into the vocabularies. There must be some optimal way to build vocabularies for TI-SPEECH so that an operator can make as few vocabulary switches as possible to fill out JAMPS message templates. Our thesis indicates that each vocabulary may need certain commands for error correction and cursor movement (i. e., DELETE, TAB BACK, TAB FORWARD, and DOWN).

We built three vocabularies for our grammar and generally put similar commands/words in certain vocabularies. For instance, we placed all alphanumerics in one vocabulary and all JAMPS function keys in another. See Appendix A for our vocabularies. Someone could definitely improve upon our version. Perhaps one large grammar of 450 words would be better to use instead of 9 vocabularies of 50 words each. But this is a technical problem--a function of signal processing within voice recognition technology. The time for computer recognition and response to voice commands would likely increase if there were just one large grammar.

A final thought to consider for operating procedures is keeping track of which vocabulary is active. Sometimes we were unsure which vocabulary was active, and therefore we lost time in filling out message templates. It was also difficult to remember which phrases were in which vocabularies, so we had the vocabularies printed out. A small display line on the computer video terminal to show the user the name of the active vocabulary would be a very helpful tool. In any case, our test grammar offers a good starting point.

2. <u>Using TI-SPEECH in a Zenith Z-150</u>

JAMPS is meant to run on microcomputers. TAFIG and MITRE have just upgraded the software so it could run on IBM microcomputers or IBM compatibles using MS-DOS as the operating system. The TI-SPEECH board is supposed to be IBM compatible. Therefore a logical experiment would be to see if TI-SPEECH will operate on a Zenith Z-150. The Z-150 is a desktop microcomputer prevalent in the Air Force and Navy. It is TEMPEST cleared and IBM compatible.

3. <u>TEMPEST and Cost Considerations</u>

To our knowledge, the TIPPC and TI-SPEECH are not TEMPEST approved so they may not be appropriate to prepare classified messages. This would have to be checked. Also the costs involved in using TI-SPEECH with a TIPPC, a current IBM, or an IBM compatible need to be examined before any decision is made to use ASR with JAMPS.

4. Collecting and Analyzing Data for the MOP

After the ASR grammar has been improved, then someone needs to test our MOP to prove or disprove that JAMPS with voice recognition saves time over JAMPS with just keyboard entry. Data would have to be collected at an exercise or under some sort of controlled environment. This research would entail experimental design and set-up in addition to data analysis.

5. Investigating Other ASR Systems

We explained why we chose the TIPPC, with TI-SPEECH installed, for use with JAMPS. We had only four ASR systems to choose from at the Naval Postgraduate School. But there are other ASR systems available for possible interfacing with JAMPS. These could be researched along the lines of our thesis and recommended research topics.

APPENDIX A

GRAMMAR FOR THE TI-SPEECH SYSTEM

The grammar we built for the TI-SPEECH is shown below in the three different vocabularies. Remember that TI-SPEECH allows up to 450 words (9 vocabularies with a maximum of 50 words each). The left hand column shows the phrase you would say; the right hand column is the command that would be input into the computer.

Pronounce the words in the left column as they are written. Make sure you always say the words the same way. As an example, let's use the phrase "552AWACS", which is a military unit. A suggestion in pronouncing this phrase is to say it the way you normally do (i.e., "Five fifty second ay wax"). Also, when you see spaces between the letters of the words/phrases in the left column, sound out each letter separately. An example is "T A C C."

* If you use the month of November, there will be a conflict between the display of "NOV" and the letter "N." A suggestion is to use a different phonetic for "N" (i.e., "NICKEL").

VOCABULARY #1 - FUNCTION

PHRASE	KEYSTROKE
start menu	<F1>
cat select	<F2>
compose	<F3>
enter argument	<F4>
help	<F5>
scroll alternate	<F6>

clear field	<F7>
repeat field	<F8>
identify KDS	<F9>
insert KDS	<F10>
new save name	<CTRL F1>
review message	<CTRL F2>
recall message	<CTRL F3>
print message	<CTRL F4>
conversation mask	<CTRL F5>
validate message	<CTRL F6>
save message	<CTRL F7>
transmit	<CTRL F8>
clear identify	<CTRL F9>
delete message	<CTRL F10>
down	<DOWN ARROW>
up	<UP ARROW>
right	<RIGHT ARROW>
left	<LEFT ARROW>
home	<HOME>
tab forward	<TAB>
tab back	<SHF TAB>
return	<RETURN>
space	<SPACE>
alphanumerics	<SWITCH TO ALPHANUM>
jamps three	<SWITCH TO JAMPS3>
delete	<BACK SPACE>
down five . .	<DOWN ARROW><DOWN ARROW><DOWN ARROW><DOWN ARROW><DOWN ARROW>

VOCABULARY #2 - ALPHANUM

PHRASE	KEYSTROKE
zero	0
one	1
two	2
three	3
four	4
five	5
six	6
seven	7
eight	8
nine	9
alpha	A
bravo	B
charlie	C
delta	D
echo	E
foxtrot	F
golf	G
hotel	H
india	I
juliet	J
kilo	K
lima	L
mike	M
november	N
oscar	O
papa	P
quebec	Q
romeo	R
sierra	S
tango	T
uniform	U

victor	V
whiskey	W
x ray	X
yankee	Y
zulu	Z
delete	<BACK SPACE>
jamps three	<SWITCH TO JAMPS3>
return	<RETURN>
space	<SPACE>
cat select	<F2>
down	<DOWN ARROW>
compose	<F3>
function keys	<SWITCH TO FUNCTION>
slant bar	/
dash	-
tab back	<SHF TAB>

VOCABULARY #3 - JAMPS3

PHRASE	KEYSTROKE
unclassified	UNCLAS
confidential	C O N F I D E N T I A L
secret	S E C R E T
efto	E F T O
airlift	AIRLF
aerial refueling	AR
close air support (or CAS)	CAS
combat air patrol (or CAP)	CAP
reconnaissance (or RECCE)	REC
19th Air Refueling Wing	19AREFW
33rd Tactical Fighter Wing	33TFW
37th Tactical Fighter Wing	37TFW
552nd AWACS	552AWACS
Commander J T F	COMJTF
T A C C	TACC
north	N
south	S
east	E
west	W
january	JAN
february	FEB
march	MAR
dash	-
space	<SPACE>
pasep	PASEP
notal	NOTAL
bold eagle 85	BOLD EAGLE 85
umpires only	UMPIRES ONLY
alphanumerics	<SWITCH TO ALPHANUM>
delete	<BACK SPACE>
return	<RETURN>

function keys . . . . <SWITCH TO FUNCTION>

g k 86 . . . . GALLANT KNIGHT-86

tab back . . . . <SHF TAB>

APPENDIX B

SPEECH RECOGNITION STUDIES AT THE NPS

The following list contains the available studies at this school. The list was compiled by

Gary K. Poock
Code 55PK
Naval Postgraduate School
Monterey, California 93943
(Phone: 408-646-2636).

These studies may be useful for any future research related to our thesis topic.

AD#

Poock, G. K., A Longitudinal Study of Five Year Old Speech Reference Patterns, Journal of the American Voice Input/Output Society (in press).

Salfer, Deborah L., Voice Automation of Ship Control, Masters Thesis, Naval Postgraduate School, Monterey, California, September 1985.

Martin, B. J. and Poock, G. K., An Initial Applied Look at Stress and Voice Recognition, Journal of the American Voice Input/Output Society, Vol. 1, No. 1, 24-33, June 1984.

#A142554 Poock, G. K. and Martin, B. J., Effects of Emotional and Perceptual-Motor Stress on a Voice Recognition Systems Accuracy: An Applied Investigation, Naval Postgraduate School Report #NPS-55-84-002, February 1984

Lombardo, J. P., Using Continuous Voice Recognition Technology as Input Medium to the Naval Warfare Interactive Support System (NWISS), Masters Thesis, Naval Postgraduate School, Monterey, California, April 1984.

#A130155 Poock, G. K. and Martin, B. J., Voice Recognition Performance with Naive Versus Practiced Users, Naval Postgraduate School Report #NPS-55-83-016, June 1983.

Poock, G. K. and Roland, E. F., Simulated TACFIRE Input Procedure for Use with Voice Data Entry, Naval Postgraduate School Report #NPS-55-83-012PR, April 1983.

#A129951 Poock, G. K., Roland, E. F. and Schwalm, N. D., Wearing Army Gas Masks While Talking to a Voice Recognition System, Naval Postgraduate School Report #NPS-55-83-005, March 1983.

#A128559 French, B. A., Some Effects of Stress on Users of a Voice Recognition System: A Preliminary Inquiry, Masters Thesis, Naval Postgraduate School, Monterey, California, March 1983.

#A128546 Yellen, H. W., *A Preliminary Analysis of Human Factors Affecting the Recognition Accuracy of a Discrete Word Recognizer for C3 Systems*, Masters Thesis, Naval Postgraduate School, Monterey, California, March 1983.

Poock, G. K., Roland, E. F. and Sparks, P. W., *Final Summary: Voice Recognition/Input Issues for TACFIRE*, Naval Postgraduate School Report #NPS-55-83-017PR, March 1983.

#A127223 Poock, G. K., Martin, B. J. and Roland, E. F., *The Effect of Feedback to Users of Voice Recognition Equipment*, Naval Postgraduate School Report #NPS-55-83-003, February 1983.

#A129975 Poock, G. K. and Roland, E. F., *Voice Recognition Vocabulary Lists for the Army's TACFIRE System*, Naval Postgraduate School Report #NPS-55-83-001, January 1983.

#A129938 Poock, G. K., Schwalm, N. D., Martin, B. J. and Roland, E. F., *Trying for Speaker Independence in the Use of Speaker Dependent Voice Recognition Equipment*, Naval Postgraduate School Report #NPS-55-82-032, December 1982.

#A129930 Poock, G. K. and Roland, E. F., *Voice Recognition Accuracy: What is Acceptable?*, Naval Postgraduate School Report #NPS-55-82-030, November 1982.

#A122470 Poock, G. K., Schwalm, N. D. and Roland, E. F., *Use of Voice Recognition Equipment with Stenographer Masks*, Naval Postgraduate School Report #NPS-55-82-028, October 1982.

Poock, G. K., Schwalm, N. D. and Roland, E. F., *Wearing Protective Masks: Effects on Voice Recognition Performance*, Proceedings of the Voice Data Entry Systems Applications Conference '82, held in San Mateo, California and sponsored by Lockheed Missiles and Space Co., September 1982.

Poock, G. K. and Roland, E. F., *Voice Recognition Input to the Army's Tactical Fire Direction System*, ONR-MIT conference on C3, Monterey, California, August 1982.

#A115757 Malarkey, T., *An Investigation of the Application of Voice Input/Output Technology in the COINS Network Control Center*, Masters Thesis, Naval Postgraduate School, Monterey, California, March 1982.

#A115744 Reuss, J., *Investigation into Air Launch Cruise Missile Flight Information Loading and Display Techniques During Flex Targeting Procedures*, Masters Thesis, Naval Postgraduate School, Monterey, California, March 1982.

#A115735 Calcaterra, F., *Application of Artificial Intelligence in Voice Recognition Systems in Microcomputers*, Masters Thesis, Naval Postgraduate School, Monterey, California, March 1982.

Poock, G. K., *To Train Randomly or All at Once...That is the Question*, Proceedings of the Voice Data Entry Systems Applications Conference, sponsored by Lockheed Missiles and Space Co., Santa Clara, California, October 1981.

#A107442 Armstrong, J. W. and Poock, G. K., Effect of Task Duration on Voice Recognition System Performance, Naval Postgraduate School Report #NPS-55-81-017, September 1981.

#A107477 Armstrong, J. W. and Poock, G. K., Effect of Operator Mental Loading on Voice Recognition System Performance, Naval Postgraduate School Report #NPS-55-81-016, August 1981.

#A102208 Poock, G. K., A Longitudinal Study of Computer Voice Recognition Performance and Vocabulary Size, Naval Postgraduate School Report #NPS-55-81-013, June 1981.

#A101823 Jay, G. T., An Experiment in Voice Data Entry for Imagery Intelligence Reporting, Masters Thesis, Naval Postgraduate School, Monterey, California, 1981.

#A105568 Taggart, J. L. and Wolfe, C. D., Speech Recognition as an Input Medium for Preflight in the P3C Aircraft, Masters Thesis, Naval Postgraduate School, Monterey, California, 1981.

#A102316 McSorley, W. J., Using Voice Recognition Equipment to Run the Warfare Environmental Simulator (WES), Masters Thesis, Naval Postgraduate School, Monterey, California, 1981.

#A103280 Neil, D. E. and Andreason, T., Examination of Voice Recognition System to Function in a Bilingual Mode, Naval Postgraduate School Report #NPS-55-81-003, February 1981.

#A102371 Batchellor, M. P., Investigation of Parameters Affecting Voice Recognition Systems in C3 Systems, Masters Thesis, Naval Postgraduate School, Monterey, California, 1981.

C027338L Bragaw, P. H., Investigation of Voice Input for Constructing Joint Chiefs of Staff Emergency Action Messages, Masters Thesis, Naval Postgraduate School, Monterey, California, 1981. (Classified)

#A106138 Elster, R., The Effects of Certain Background Noise on the Performance of a Voice Recognition System, Naval Postgraduate School Report #NPS-55-80-010, September 1980.

#A093557 Armstrong, J. W., The Effects of Concurrent Motor Tasking on Performance of a Voice Recognition System, Masters Thesis, Naval Postgraduate School, Monterey, California, 1980.

#A091055 Poock, G. K., Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network, Naval Postgraduate School Report #NPS-55-80-016, Monterey, CA, 1980.

Anyone affiliated with a research and development activity within the U. S. Government or its associated contractors, subcontractors, or grantees, and under current U. S. Government contract, can order reports by AD# from

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22304.

All others not within these categories may order from

United States Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Phone: (703) 487-4650/4660 TELEX: 89-9405.

APPENDIX C

USER'S GUIDE FOR USING TI-SPEECH

This user's guide will provide the basic guidelines and commands an individual will need to bring up the TI-SPEECH on the TI Portable Professional Computer and use it as an input device for his computer. We will operate a basic TIPPC system with all the hardware and software (described in Chapter 4) installed. If the software has not been installed, the user will have to refer to the appropriate manuals for the necessary procedures.

Chapter 4 addressed several guidelines to follow when implementing the speech recognition system. Keep these guidelines in mind as you use the system. Texas Instruments' guidance is that once the voiceprints have been created, the user should test his vocabulary and then refine those voiceprints in which the matches were low or poor (less than a match value of 5). We had greater success in recognition when we refined all the voiceprints before executing the test option. We used three passes during refinement. Further study should determine the optimal number of refinement passes to achieve the best possible recognition.

Following each explanatory paragraph are the commands or instructions to execute the necessary procedures. When you see the command <RETURN>, it means to press the RETURN key on the lower right portion of the main keyboard area. <ENTER> means to press the ENTER key on the lower right hand side of the numeric keypad. <Fnn> indicates the function key corresponding to the number, nn, is to be pressed. Commands inside the quotation marks are commands the operator must enter from the keyboard. Do not enter the quotes.

Cursor instructions will also appear inside brackets with the command segment to be selected in single quotation marks. A sample command would look like this:

"EMULATE" <RETURN>

This means the operator should type the word, emulate, and then press the RETURN key to execute the command. A sample instruction would look like this:

<MOVE CURSOR TO 'ADD AN ENTRY'> <RETURN>

This instruction means the operator should use the cursor to highlight the segment "ADD AN ENTRY" and then press the RETURN key.

When in the Vocabulary Manager, the user should remember the RETURN key is used to select a command segment and build a command for execution. Once the command is built, press the ENTER key to execute the command. There are many instances in which the RETURN key will result in a command being executed. The user will need to be alert to the prompts in this guide and those provided by TI.

At every point TI has provided a help function for the user. You can read it by pressing the appropriate function key as shown in the function key area of each menu. The help function provides additional information on the command segment that is highlighted at the time. This information can clear up many questions a user may have during use of the VM. You can use the help function any time without losing your place in the VM.

<u>Steps to Enter the Vocabulary Manager</u>.

1) Turn on the computer. The power switch is located on the top center of the TIPPC near the back edge.

2) The computer will perform a series of power-up self checks and display some routine messages. Finally the system will show a date and ask the user to input the current date. Enter the date using all numbers. For example, you would enter March 8, 1986, as 3-8-86. Or

the user can simply bypass the date by pressing the RETURN key.

<RETURN>

3) Next the system will ask for the correct time. You should enter time using the 24 hour clock. Two o'clock in the afternoon would be entered as 14:00. As with the date the user can bypass the time simply by pressing the RETURN key.

<RETURN>

4) The system will now automatically execute a series of commands and bring up the NaturalLink program. The user will see a NaturalLink introductory screen and then the primary NaturalLink menu. The user should now use the function key to exit the NaturalLink program.

<F12>

5) The user will now see the DOS prompt, E>. This indicates the default drive is the Winchester hard disk. This means the computer will look for the executable programs on the hard disk. Whenever the prompt appears, the computer is waiting for the next command. First the user must enter the speech subdirectory. This is where the Vocabulary Manager and entered vocabularies reside. Note the reverse slash stroke in the command.

"CD \TISPEECH" <RETURN>

6) Next the device service routines (DSR) must be installed as discussed in Chapter 4. This is done by typing TISPEECH and pressing the RETURN key. While the routines are being installed, a message will appear. When the DOS prompt reappears, go on to step seven (7).

"TISPEECH" <RETURN>

7) At last the user is ready to activate the Vocabulary Manager. Type in VM and hit the RETURN key.

"VM" <RETURN>

8) The user will now see the Vocabulary Manager logo. At this point the VM asks the user to enter his name. By entering his name the user establishes the subdirectory in which his voiceprints will be stored.

<ENTER YOUR NAME> <RETURN>

Steps for Entering/Training a New Vocabulary

The user will now see the Vocabulary Manager's main menu. Using this menu the user will be able to execute all of the VM functions. To enter and train a new vocabulary for use, follow these steps:

1) Note on the main menu the first item in the command section is highlighted. This item is "set up and

train a vocabulary". Select this item using the RETURN key.

<RETURN>

2) Now the highlight has moved to the first item in the vocabulary listing. Using the cursor control keys, move the highlight until the "<other vocabulary>" is highlighted and select this segment with the RETURN key. (You can also hit the HOME key, which is in the middle of the five cursor-key cluster, to bring you to "<other vocabulary>," which is at the bottom of the vocabulary listing).

<MOVE CURSOR TO '<OTHER VOCABULARY>'> <RETURN>

3) Now a pop-up window will appear asking for the name of this vocabulary. Enter the name chosen for the new vocabulary and press the RETURN key.

<ENTER VOCABULARY NAME> <RETURN>

4) Note the command has now been completed and appears in the command window. The user is prompted to press the run(ENTER) key to execute the command.

<ENTER>

5) This brings up a new menu for the user. This is the "set up and train" menu and is used in the creation and training of new vocabularies or the training of already entered vocabularies. The item "add an entry" is highlighted; select this item by pressing the RETURN key.

<RETURN>

6) A pop-up window will appear prompting the user to enter the phrase for the command he wants to execute. Enter the phrase and press the RETURN key.

<PHRASE> <RETURN>

7) A second pop-up window will appear requesting the user to input the keystrokes corresponding to the phrase and command to be executed. The user should keep in mind that each key pressed will appear in the pop-up window. If an error is made it will be necessary to use the "modify keystrokes" command after the entire set of keystrokes has been entered. Once the user enters the keystrokes, he will have to press and hold the CONTROL (CTRL) key (found at the left side of the main keyboard area) and press the RETURN key at the same time to indicate he has completed entry of the keystrokes.

<ENTER KEYSTROKES>
<CTRL> <RETURN> simultaneously

8) This will complete the entry of both phrase and keystrokes. Note the "set up and train vocabulary" menu has returned with the "add an entry" item highlighted. The user should continue to execute steps five (5), six (6), and seven (7) until he has entered his entire vocabulary. When this has been completed, go on to step nine (9).

9) Use the cursor control keys to highlight the "create voiceprints" command and press the RETURN key.

<MOVE CURSOR TO 'CREATE VOICEPRINTS'> <RETURN>

10) A pop-up window will appear with three options. The option "all of them" is already highlighted. Press the RETURN key.

<RETURN>

11) The user will be prompted to insure the input device is turned on. When the user is ready to begin, he should press the RETURN key. The computer now prompts you to say each word as it appears in another pop-up window.

<RETURN>

12) When the user has created his voiceprints, he should refine them. Move the highlight to the "refine voiceprints" item and press the RETURN key.

<MOVE CURSOR TO 'REFINE VOICEPRINTS'> <RETURN>

13) A pop-up window will appear asking the user which voiceprints he wishes to refine. The user should select the option "all of them."

<RETURN>

14) A window will now appear asking the user how many times he wishes to refine each voiceprint. The user should enter the number three (3) and press the RETURN key.

"3" <RETURN>

15) The user is now prompted to insure the input device is turned on and in place. When the user is ready to execute the refinement process, he should press the RETURN key.

<RETURN>

16) The user will now be prompted to say the phrase three times. The number increments each time the phrase is recognized. When you have said the phrase three times, the next phrase will appear for refinement until the entire vocabulary has been refined.

17) When the refinement is complete, the "set up and train" menu will once again appear. The user should now highlight and select the "test voiceprints" command.

<MOVE CURSOR TO 'TEST VOICEPRINTS'> <RETURN>

18) All of the phrases will appear on the screen behind a pop-up window which prompts the user to insure the input device is activated. When he is ready to begin the test, the user should press the RETURN key. Each phrase is spoken until the entire vocabulary has been tested. The user will see the match window, discussed on page 53 in Chapter 4, for each phrase spoken.

<RETURN>

19) Once the user completes the test, he should press the ENTER key.

<ENTER>

20) If the user was not satisfied with all the matches he achieved, he can refine those particular voiceprints using the "refine voiceprints" command. If the user is satisfied, he should press the F10 key to exit the "set up and train" mode and return to the main menu. The user must insure he saves his voiceprints by pressing the RETURN key when asked.

<RETURN><F10> <u>or</u> <CURSOR 'REFINE VOICEPRINTS'><RETURN>

<u>Summary of Steps for VM Commands</u>

A general summary of the steps above is provided here for the user's benefit. These steps draw together the main sequence of actions when entering and training a new vocabulary.

1) Select the "set up and train" command.
2) Enter the new vocabulary name.
3) Add each new entry.
4) Create the voiceprints for the new vocabulary.
5) Refine each voiceprint three times.
6) Test your voiceprints to determine level of match.
7) Refine any voiceprints you feel are not adequate.
8) Exit the "set up and train" menu. Be sure to save your voiceprints when prompted.

<u>Training an Already-Created Vocabulary</u>

To train a vocabulary that has already been entered, the user should use the following steps. The main menu should be displayed.

1) Highlight the item "set up and train a vocabulary" and press the RETURN key.

<CURSOR TO 'SET UP AND TRAIN VOCABULARY'> <RETURN>

2) Note the highlight has now moved to the vocabulary listing. Using the cursor control keys, highlight the name of the vocabulary you want to train. Select this vocabulary by pressing the RETURN key.

<SELECT VOCABULARY NAME> <RETURN>

3) The vocabulary will appear in the command window, and the user will be prompted to press the run(ENTER) key.

<ENTER>

4) The "set up and train" menu will now appear. The user should execute the steps outlined in the previous section starting with step nine (9).

Adding a Voice-Activated Vocabulary Switch

As discussed in Chapter 4, there are two methods of switching from one vocabulary to another vocabulary (changing the active vocabulary). This can be done by using the main menu or having the computer execute the change with a voice command. We considered using a voice activated switch to be faster. To execute a voice activated switch, an appropriate phrase and keystrokes must first be entered. To do this the user should follow these steps:

1) The user sh[illegible] perform steps one (1) and two (2) of the secti[illegible] entering/training a new vocabulary. Instead [illegible]ing "<other vocabulary>" in step two (2), se[illegible] name of the vocabulary that needs the vocabular[illegible]ch added.

2) When the "set up and train menu" appears, execute steps five (5) through seven (7) in the entering/ training a new vocabulary section. The phrase to be entered is the phrase that, when spoken, will cause the computer to execute the switch. We used the name of the vocabulary we wanted to switch to as our phrase (i. e., switch word). The keystrokes at this point are optional. If you want the computer to execute some command before it makes the switch, then keystrokes should be entered. If you only want the computer to change the active vocabulary, then no keystrokes need to be entered.

3) You will now have entered the switch phrase and keystrokes, if desired. The "set up and train" menu should now be displayed. Using the cursor control keys, highlight the "modify keystrokes" command and select it by pressing the RETURN key.

   <MOVE CURSOR TO 'MODIFY KEYSTROKES'> <RETURN>

4) The vocabulary will appear. Using the cursor control keys, highlight the vocabulary switch you have just entered and select it with the RETURN key.

   <MOVE CURSOR TO THE APPROPRIATE ENTRY> <RETURN>

5) A new menu, "modify keystrokes", will now appear. Using the cursor control keys, highlight the command "add a vocabulary switch" and select it by pressing the RETURN key.

   <MOVE CURSOR TO 'ADD A VOCABULARY SWITCH'> <RETURN>

6) All the available vocabularies will now appear in a pop-up window. Two other entries are also available. First is the "<previous vocabulary>." Use of this choice will take the user back to the vocabulary that was active immediately before the one now active. The

other choice is "<other vocabulary>." This lets the user enter the name of a vocabulary that he plans to enter. Select the vocabulary name or command that is desired and press the RETURN key.

<MOVE CURSOR TO APPROPRIATE CHOICE> <RETURN>

7) The user has now completed the entry of the vocabulary switch and should press the F10 key to exit the "modify keystrokes" menu.

<F10>

8) The user will now see the "set up and train vocabulary." You should now perform steps 9 through 20 of the section, entering/training a new vocabulary, to create the needed voiceprints. Make sure the vocabulary is ready for use.

As you insert switch words into a vocabulary, make sure there is a command for each vocabulary that you might wish to link together. If you are planning to use nine vocabularies at one time, then consider putting a switch word for each additional vocabulary in one master vocabulary. One switch word in each of the eight additional vocabularies is all that will be necessary to return to the master vocabulary and then to another vocabulary. This way valuable vocabulary space is not taken up by a large number of switch phrases.

Starting Voice Recognition

Once all the appropriate vocabularies have been entered and trained, the user is ready to turn on the voice recognition and implement the chosen application accordingly. To start the voice recognition, follow these steps:

1) The main menu should be displayed. Using the cursor control keys, highlight the "start recognition of vocabulary" and hit the RETURN key.

<CURSOR TO 'START RECOGNITION OF VOCABULARY'> <RETURN>

2) The first item in the vocabulary listing is highlighted. Using the cursor control keys, move through the list and highlight the desired vocabulary. Press the RETURN key.

<SELECT DESIRED VOCABULARY> <RETURN>

3) If the user has only one vocabulary to use, he should now skip to step 8. If the user wishes to link several vocabularies together, then follow the next steps.

4) Once the first vocabulary is selected, the user will note another segment becomes available to be chosen. The selection, "and vocabulary," is highlighted already. Hit the RETURN key.

<RETURN>

5) After you select the "and vocabulary" segment, the first item in the vocabulary listing is once again highlighted. Using the cursor control key, highlight the next vocabulary to be used and press the RETURN key.

<SELECT NEXT VOCABULARY> <RETURN>

6) The user will now note the "and vocabulary" has again been highlighted. An additional choice has appeared. This is the "set the first active vocabulary to" segment. If the the user wishes to link more vocabularies together, he should repeat steps four (4) and five (5) above until all have been selected. Remember the maximum number that can linked is nine (9). Once all the vocabularies have been selected you must set the active vocabulary. This is the vocabulary which the ASR system will first recognize when enabled. To do this, select the "set first active vocabulary to" segment with the cursor control and press the RETURN key.

<SELECT 'SET FIRST ACTIVE VOCABULARY TO'> <RETURN>

7) When you choose the "set first active vocabulary to" segment, all the previously selected vocabularies will appear in a pop-up window. The user should select the vocabulary he wishes to be active with the cursor control key and press the RETURN key.

<SELECT DESIRED ACTIVE VOCABULARY> <RETURN>

8) The command has now been completed and is ready for execution. When the user is ready, he should press the ENTER key. The system will blank the screen and a brief message will appear while the vocabularies are loaded into memory.

<ENTER>

9) When all the vocabularies are in memory and the ASR system is activated, the main VM menu will be redisplayed. The user should exit VM by pressing the F10 key.

<F10>

10) The DOS prompt (E>) will appear indicating the system is ready to implement the desired application using voice or keyboard input.

<u>Using Voice Recognition With JAMPS</u>

This thesis has been concerned with the application of voice recognition to the JAMPS II.V. The JAMPS software, in a version usable by the TIPPC, now resides on one of the

TIPPC's Winchester hard disk. This software uses over five megabytes of memory. To use voice input with the JAMPS software, follow these steps:

1) The vocabularies we designed for use with JAMPS are called ALPHANUM, FUNCTION, and JAMPS3. You will find the contents of each vocabulary in Appendix A. The user must first create his voiceprints for these vocabularies by performing the steps listed in the above section, Training an Already-Created Vocabulary.

2) Once you have created the voiceprints, you must execute the steps in the Starting Voice Recognition section. It does not matter which vocabulary the user sets as active first. That choice can be better made after the user has gained some experience in using JAMPS.

3) When the DOS prompt appears, the user must tell the TIPPC to emulate an IBM Personal Computer. Do so by executing the EMULATE program available in the operating system provided by TI. This program must be executed for JAMPS to run on the TIPPC.

   "EMULATE" <RETURN>

4) A message will appear telling the user the emulator has been run. Then the DOS prompt will be displayed. The user must now enter the subdirectories where the JAMPS programs reside. This is done by entering the following command. Note the reverse slashes.

   "CD \OED\JAMPS" <RETURN>

5) The DOS prompt will again appear. This time enter JAMPS and press the RETURN key. This will bring up the initial JAMPS menu, as seen on page 21 in Chapter 2. JAMPS is now available and ready for user manipulation using either voice or keyboard input.

   "JAMPS" <RETURN>

Guidelines for Filling in a JAMPS Message Template with Voice Recognition

These are guidelines for manipulating a JAMPS message template with voice input commands. As an example, we will use the B704 mask mode template, as shown in Figure 2.3 You will find the vocabularies of available words for input in Appendix A. Please note that this is not a complete listing of all commands or input words to use with JAMPS. We will cover a few of the input commands in our grammar. For complete information on all JAMPS commands, see the JAMPS II.V Operator's Manual.

Each paragraph will explain the sequence of voice commands for input. Following the paragraph will be the appropriate voice commands. Voice commands will appear in quotation marks.

1) Ensure the ALPHANUM vocabulary is active, and the other two vocabularies, JAMPS3 and FUNCTION, are available. If you have completed the steps in the previous section, then you should see the initial JAMPS main menu. Select the category, AIR OPERATIONS, by saying "2." Then switch to the FUNCTION vocabulary by saying "FUNCTION KEYS." Execute your selection by saying "CAT SELECT."

   "2" "FUNCTION KEYS" "CAT SELECT"

2) You will now see the listing of AIR OPERATIONS message templates. Switch back to the ALPHANUM vocabulary by saying "ALPHANUMERICS." Next say the message number, B704. Switch back to the FUNCTION vocabulary and compose the message.

   "ALPHANUMERICS" "B704" "FUNCTION KEYS" "COMPOSE"

3) You will see the cursor located in the first data set on the first line. Addressing information from a previous JAMPS demonstration is already placed in the heading portion of the message. Move the cursor down to the EXER/EXERNICK set by saying "DOWN FIVE" two times and "DOWN" three times.

   "DOWN FIVE" "DOWN FIVE" "DOWN" "DOWN" "DOWN"

4) Say "TAB FORWARD" to move the cursor to the beginning of the empty data field. Note that the blanks in the data field convert to map characters, which are discussed in Chapter 2. Change to the JAMPS3 vocabulary by saying "JAMPS THREE." Enter "BOLD EAGLE 85" into the field, then say "RETURN."

   "TAB FORWARD" "JAMPS THREE" "BOLD EAGLE 85" "RETURN"

5) The user should continue to input information into the empty data fields in a similar manner by using the appropriate phrases and switch words from Appendix A. You will see a highlighted status line at the bottom of the screen at all times. This status line tells you the minimum and maximum number of characters for that field and whether or not the field is mandatory.

6) If at any time you do not know what information is required in a field, you can enter the HELP mode. Ensure you are in the FUNCTION vocabulary before saying "HELP." As we already stated in Chapter 2, the user can also enter the CONVERSATION MASK mode in order to fill out data fields. In this mode you can fill out each data field with the aid of HELP screens. A HELP screen gives you the field you are located in, the allowable data entries for that field, and a definition of the field. After you enter the data for the field and say "RETURN," JAMPS will display for you the next available field in a HELP function set-up. You may continue to fill out the entire message in CONVERSATION MASK mode.

7) You should remember that our grammar is not large enough to complete the entire message. However, the grammar is sufficient to demonstrate the feasibility of using voice recognition as an input mode to JAMPS.

8) Once you have filled out the message template, you have several options on what to do next. The FUNCTION vocabulary has the phrases to review, print, save, validate, delete, or transmit the message. Refer to the JAMPS II.V Operator's Manual for a full explanation of these functions.

9) To exit JAMPS, ensure the FUNCTION vocabulary is active. Say "START MENU," and you will get the original JAMPS main menu. Due to shortcomings in the interface between the JAMPS software and the TIPPC, there is no graceful way to exit JAMPS. Entering "10" "CAT SELECT" will not work off of the main menu. Therefore, you must reboot the system by pressing the CTRL, ALT, and DEL keys simultaneously. This will bring you back to computer system start-up, as outlined in the first section of this appendix.

   "FUNCTION KEYS" "START MENU"
   <CTRL> <ALT> <DEL> simultaneously

In this user's guide we provided the basic guidelines for using the TI-SPEECH Vocabulary Manager and JAMPS. More detailed instructions for the VM commands we provided are in the TI-SPEECH Vocabulary Manager Guide. You will also find instructions on additional VM commands available in this same reference. For information on using the JAMPS software, refer to the JAMPS II.V Operator's Manual.

LIST OF REFERENCES

1. Lawson, Joel S., Sr., "The State Variables of a Command Control System," Proceedings for Quantitative Assessment of the Utility of Command and Control Systems, Office of the Secretary of Defense with cooperation of MITRE Corporation, C3 Division. National Defense University, Washington D.C., January 1980.

2. U.S., Joint Chiefs of Staff, Publication Number 1, Department of Defense Dictionary of Military and Associated Terms. U.S. Government Printing Office, Washington D.C., 1 June 1979.

3. MITRE Corporation Technical Report 9372, "JAMPS Overview Description," by Paul R. Bigelow, August 1984.

4. Archibald, Norman E., Major General (USA), and Michelli, Thomas J., "JTC3A: Joint Tactical Command, Control and Communications Agency," SIGNAL, Journal of the Armed Forces Communications and Electronics Association, Vol 39 No 3, November 1984.

5. U.S., Joint Tactical Command, Control, and Communications Agency, "JINTACCS Quick Reference Guide," Washington D.C., January 1985.

6. Case, Randolph M., Major (USAF), and Morrison, John S., Major (USAF), "Of JINTACCS and Jabberwockies," SIGNAL, Journal of the Armed Forces Communications and Electronics Association, Vol 38 No 3, November 1983.

7. U.S., Joint Tactical Command, Control and Communications Agency, "JTC3A Major Project Summary," Fort Monmouth, New Jersey, 26 June 1985.

8. "JINTACCS User's Handbook for Joint Tactical C2 Elements, Message Text Formats (MTF)," System Development Corporation in cooperation with U.S. Readiness Command, 1 May 1985.

9. Blohm, William J., and Brown, William R., Lt Col (USAF), "Joint Tactical C3 Architecture," SIGNAL, Journal of the Armed Forces Communications and Electronics Association, Vol 40 No 3, November 1985.

10. U.S., Tactical Air Forces Interoperability Group, JINTACCS Program Office, "JAMPS II.V Operator's Manual," Langley Air Force Base, Virginia, March 1985.

11. MITRE Corporation Working Paper 24142, "JAMPS II System Specification," by Paul R. Bigelow, 23 March 1983.

12. MITRE Corporation Technical Report 8480, "JAMPS C & I System Computer Program Design Plan, by JAMPS Working Group, August 1982.

13. Jane's Military Communications 1985, sixth edition, Jane's Publishing Company Limited, London, England, 1985.

14. White, George M. "Speech Recognition: An Idea Whose Time is Coming," Byte, January, 1984.

15. Cupples, E. J., and Woodard, J. P. "Selected Military Applications of Automatic Speech Recognition Technology," SPECTRUM Vol 20. December, 1983.

16. Manson, Rick B., and Wright, Michael E., "Comparison of Continuous Speech, Discrete Speech and Keyboard Input to an Interactive Warfare Simulation in Various C3 Environments." Thesis done at Naval Postgraduate School, Monterey, California, March 1985.

17. Texas Instruments, Incorporated, Operating Instructions - Texas Instruments Portable Professional Computer, TI part no. 2238689-0001, Austin, Texas, 1983.

18. Texas Instruments, Incorporated, Technical Reference Manual, TI part no. 2223216-001, Austin, Texas, 1983.

19. Texas Instruments, Incorporated, TI-SPEECH Application Toolkit Guide, TI part no. 2232384-0001, Austin, Texas, 1985.

20. Texas Instruments, Incorporated, TI-SPEECH Vocabulary Manager Guide, TI part no. 2223202-0001, Austin, Texas, 1985.

21. Military Operations Research Society (MORS) C2 MOE Workshop, Command and Control Evaluation Workshop, ed. Dr. Ricki Sweet, Dr. Morton Metersky, and Dr. Michael Sovereign (unpublished workshop report, Naval Postgraduate School, Monterey, California, November 1985).

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Prof. Gary K. Poock, Code 55PK Naval Postgraduate School Monterey, California 93943-5000	2
4.	TAFIG/II (Attn: Colonel Jack Ezzelle) Langley AFB, Virginia 23665-5000	1
5.	TAFIG/IISJA (Attn: Major L. P. Tucker) Bldg 27, 6th Street Langley AFB, Virginia 23665-0716	1
6.	Superintendent, Code 74 (Attn: Dr. Mike Sovereign) Naval Postgraduate School Monterey, California 93943-5000	6
7.	HQ TAC/INYW (Attn: Captain Leo B. Kotowski) Langley AFB, Virginia 23665-5000	2
8.	JTC3A (DSR&E) Washington Field Activity (Attn: Captain Frank Prautzch) JINTACCS Program Office Fern Street Washington, D. C. 20301-3160	1
9.	AFIT/CISK (Attn: Major Harlan) Wright-Patterson AFB, Ohio 45433-6583	1
10.	U.S. Army Electronics Technology and Devices Laboratory (Attn: SLCET-DA, Captain Earl T. Hill) Fort Monmouth, New Jersey 07703-5302	1
11.	Mike Findlay Rt 365 Hinckley, New York 13352	1
12.	Bill Caldwell 3148 Mimi Court Marina, California 93933	1
13.	TEXAS INSTRUMENTS (Attn: Roy Pursley) P.O. Box 2909, MS 2081 Austin, Texas 78769	1

14. Commander 1
Naval Ocean Systems Center
(Attn: Mike Mudarian)
San Diego, California 92152

15. Major Thomas J. Brown 1
Code 39
Naval Postgraduate School
Monterey, California 93943-5000

16. Superintendent 1
Curriculum Office, Code 39
Naval Postgraduate School
Monterey, California 93943-5000

17. RADC/IRAA 3
(Attn: Lt. Stockton)
Griffiss AFB, New York 13441-5700

18. OJCS/C3S 1
Washington, D.C. 20301

19. U.S. Army CECOM 1
AMSEL-SEI-F
(Attn: Dr. Israel Mayk)
Fort Monmouth, New Jersey 07703

END