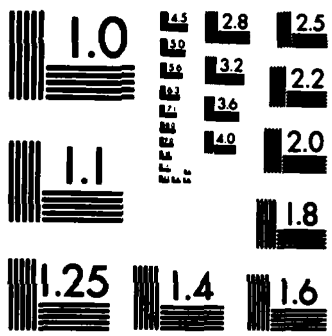
AD-A143 036 DEVELOPMENT OF A COMPUTER-ASSISTED SIMULATION OF TACTICAL VOICE COMMUNICA..(U) HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA D L HANNAMAN ET AL. DEC 83 1/1

UNCLASSIFIED HUMRRO-FR-TRD(VA)-83-7 MDA903-82-C-0511 F/G 17/2 NL



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

Development of a Computer-Assisted Simulation of Tactical Voice Communications

Phase I — Conceptual Design

David L. Hannaman
Human Resources Research Organization

Paul R. Michaelis
Texas Instruments, Incorporated

Randall M. Chambers
U.S. Army Research Institute for the
Behavioral and Social Sciences

HUMAN RESOURCES RESEARCH ORGANIZATION
1100 South Washington Street • Alexandria, Virginia 22314-4499

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A143 036
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Development of a Computer-Assisted Simulation of Tactical Voice Communications: Phase I - Conceptual Design	5. TYPE OF REPORT & PERIOD COVERED Final Report Sept '82 - Dec '83
	6. PERFORMING ORG. REPORT NUMBER Final Report 83-7
7. AUTHOR(s) David L. Hannaman (HumRRO) Paul R. Michaelis (Texas Instruments) Randall M. Chambers (ARI)	8. CONTRACT OR GRANT NUMBER(s) MDA903-82-C-0511
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Ave., Alexandria, VA 22333	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q162717A790
11. CONTROLLING OFFICE NAME AND ADDRESS ---	12. REPORT DATE December 1983
	13. NUMBER OF PAGES 94
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) ---	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

---

18. SUPPLEMENTARY NOTES

---

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Speech Synthesis	Tactical Communications
Voice Synthesis	Soldier/Machine Interface
Speech Recognition	Battle Simulation
Voice Recognition	Artificial Intelligence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The dramatic introduction of technology onto the battlefield has intensified in recent years and is expected to continue at an accelerated rate. This results in increased human performance demands. Human performance is physiologically limited in terms of the rate at which a human can process information and physically perform tasks. Though there is little that can be done about the human's physiological limitations, human performance capabilities can be improved in other ways. Human performance capabilities of Army personnel can

-- continued --

20. ABSTRACT (continued)

best be improved by either raising enlistment standards (e.g., education, aptitude) and/or providing more and better training. Enlistment standards are expected to remain relatively constant during the foreseeable future. Though Army training has improved and will continue to do so, the training environment is burdened with everincreasing demands (to a large extent attributable to the proliferation of technology on the battlefield) while simultaneously experiencing diminishing resources (e.g., time and money). As a result, human performance capabilities in the Army are also expected to remain relatively constant.

Given a relatively constant soldier performance capability and increased complexity of the war machines which the soldier must operate, a serious problem surfaces. Today's soldier is expected to operate complex, technologically sophisticated systems upon which victory on the battlefield is highly dependent. One way to lessen the gap between soldier performance capabilities and technologically advanced systems is to improve the soldier/machine interface (SMI).

The SMI can be improved in a variety of ways including environmental/job design, artificial intelligence, decision aids and voice recognition/synthesis. The objective of this research is to advance the application of evolving speech technology to Army tactical operational and training systems. To achieve this objective, a conceptual design was developed for a research vehicle designed to assess the effects of a voice input/output (I/O) soldier machine interface (SMI) on soldier performance. To accomplish this, a review of the state-of-the-art of voice technology was performed and its potential benefits to tactical operational and training systems' SMIs determined. A definition and taxonomy of tactical voice communications were developed. A conceptual design of a computer-assisted simulation of tactical voice communications (SIMCOMM) was then developed which includes its voice interactive protocols, i.e., speech synthesis and recognition requirements, and a specification of its hardware configuration.

Accomplishing the activities covered in this report constitutes completion of Phase I of this research effort. Phase II will encompass the actual building and testing of the SIMCOMM.

Final
Report
83-7

HumRRO
FR-TRD(VA)-83-7

HumRRO

Development of a Computer-Assisted Simulation of Tactical Voice Communications

Phase I — Conceptual Design

David L. Hannaman
Human Resources Research Organization

Paul R. Michaelis
Texas Instruments, Incorporated

Randall M. Chambers
U.S. Army Research Institute for the
Behavioral and Social Sciences

HUMAN RESOURCES RESEARCH ORGANIZATION
1100 South Washington Street • Alexandria, Virginia 22314-4499

December 1983

Prepared for:

**U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333**

DEVELOPMENT OF A COMPUTER-ASSISTED SIMULATION OF TACTICAL VOICE COMMUNICATIONS

PHASE I — CONCEPTUAL DESIGN

Brief

REQUIREMENT

The dramatic introduction of technology onto the battlefield has intensified in recent years and is expected to continue at an accelerated rate. This results in increased human performance demands. Human performance is physiologically limited in terms of the rate at which a human can process information and physically perform tasks. Though there is little that can be done about the human's physiological limitations, human performance capabilities can be improved in other ways, i.e., raising enlistment standards (e.g., education, aptitude) and/or providing more and better training. Enlistment standards are expected to remain relatively constant during the foreseeable future. Though Army training has improved and will continue to do so, the training environment is burdened with everincreasing demands (to a large extent attributable to the proliferation of technology on the battlefield) while simultaneously experiencing diminishing resources (e.g., time and money). As a result, human performance capabilities in the Army are also expected to remain relatively constant.

Given a relatively constant soldier performance capability and increased complexity of the war machines which the soldier must operate, a serious problem surfaces. Today's soldier is required to do more than merely point his rifle. He is expected to operate complex, technologically sophisticated systems upon which victory on the battlefield is highly dependent. One way to lessen the gap between soldier performance capabilities and technologically advanced systems is to improve the soldier/machine interface (SMI).

PROCEDURE

The SMI can be improved in a variety of ways including environmental/job design, artificial intelligence, decision aids and voice recognition/synthesis. The objective of this research is to advance the application of evolving speech technology to Army tactical operational and training systems. To achieve this objective, a conceptual design was developed for a research vehicle designed to assess the effects of a voice input/output (I/O) soldier machine interface (SMI) on soldier performance. To accomplish this, a review of the state-of-the-art of voice technology was performed and its potential benefits to tactical operational and training system's SMIs determined. A definition and taxonomy of tactical voice communications were developed. A conceptual design of a computer-assisted simulation of tactical voice communications (SIMCOMM) was then developed which includes its voice interactive protocols, i.e., speech synthesis and recognition requirements, and a specification of its hardware configuration.

Accomplishing the activities covered in this report constitutes completion of Phase I of this research effort. Phase II will encompass the actual building and testing of the SIMCOMM.

FINDINGS

The critical role SMIs play in military systems is identified and the potential benefits of voice I/O SMIs to both operational and training systems discussed. A survey and evaluation of currently available voice synthesis/recognition and related computer technologies is presented from which an optimal approach to developing a computer-assisted simulation of tactical voice communications is determined.

A definition and taxonomy of tactical voice communications are presented based upon a review of both formal Army and research literature. Tactical voice communications are defined in terms of units involved, communication means, communication equipment, and communication nets. Its taxonomy identifies tactical voice communications as consisting of seven components, i.e., content classifications, radio/telephone procedural terminology, objects, phonetic alphabet, numbers, jamming/interference, and background noises.

To satisfy the objectives of this research, SIMCOMM's general requirements are identified as being a standalone system, capable of demonstrating integrated technologies, and able to satisfy specific application parameters (e.g., involve voice oriented tasks, be limited in scope, and common to small unit tactical operations). Given these general requirements, "Call for Fire" was selected as the basic scenario around which SIMCOMM would be designed. SIMCOMM's conceptual design is presented in terms of its operational concept, basic hardware configuration, and interactive voice protocols. The approach to SIMCOMM's development, testing and expansion are then presented.

UTILIZATION OF FINDINGS

The SIMCOMM is a standalone, portable system which will demonstrate voice synthesis/recognition and, to a lesser degree, artificial intelligence technologies. Its associated hardware/software configuration is compatible with most mini/micro and large mainframe computers currently used in the Army. SIMCOMM's software is modular by design. Given these attributes, SIMCOMM can be transported for demonstration and research purposes, interfaced to existing tactical operational and training systems, and its speech synthesis/recognition capabilities easily expanded. Most important, SIMCOMM can be used as a research tool for investigating voice input/output soldier/machine interfaces. SIMCOMM's cost is minimal and, as such, facilitates its replication. Combined, SIMCOMM's attributes and capabilities will enable it to advance the application of voice technology in Army environments. The definition and taxonomy of tactical voice communications developed may be utilized in whole or part for further investigations or analysis of communications within a tactical, small unit environment. Areas of interest would include the effect of communications on small unit performance, assessment of tactical communication training needs, and facilitating the incorporation of voice synthesis/recognition technologies in existing or planned operational (e.g., VINT[2]) or training (e.g., battle simulations) systems.

TABLE OF CONTENTS

Page

Page

1 — BACKGROUND AND INTRODUCTION

Speech synthesis and speech recognition technologies have advanced considerably in the past few years. This report describes a research effort that is exploring Army applications for these technologies. As part of this effort, a versatile prototype voice input/output (I/O) soldier/machine interface (SMI) is being developed. This interface will be used to investigate the potential benefits of voice I/O SMIs in Army operational and training systems.

BACKGROUND

The requirement for a computer-assisted simulation of tactical voice communications stems from problems encountered with existing soldier/machine interfaces (SMI) in two application areas—operational systems and training systems. Operational systems include the actual hardware or systems used by the Army for military missions, e.g., weapon systems and vehicles. The training systems of concern address tactical training requirements, such as tactical engagement simulation and battle simulations.

The frequency and extent of technological advancement in Army operational systems present a serious dilemma. The rapid and major introduction of this technology into Army operational systems has resulted in their becoming more complex. Meanwhile, the performance capability of these systems' operators remains relatively constant. This is true for several reasons. These include human physiological or psychomotor limitations (i.e., volume and rate at which a human can process information and physically perform tasks), little change in enlistment standards (e.g., education and aptitude) during the foreseeable future, and an Army training environment characterized by everincreasing training demands and simultaneously diminishing training resources.

A dramatic change in the technological complexity of Army operational systems has been experienced over the past fifty years. This is expected to continue at an even more dramatic rate. During this same period, human performance capabilities have remained constant for the reasons previously discussed. The "gap" between technology complexity and human performance is the problem. The gap will become even greater when the soldier must survive and defeat an enemy in an AirLand Battle in the year 2000. One way to lessen this gap is to improve the soldier/machine interface.

Acute SMI problems have also been encountered in tactical training systems. A specific example of these can be found in battle simulations (BS) upon which the Army is becoming increasingly dependent for tactical training. Though BSs have been shown to improve tactical proficiency, these training techniques need a sizable, costly control team, i.e., a need to populate the systems with humans (controllers). This control requirement is defined by the very nature of the simulated environment: a complex, dynamic, and continuous flow of information is necessary for command and control of the battlefield. As a result of this "human population" requirement, a typical BS may take several hours to complete what would normally be only a few minutes of real time on an actual battlefield. Players and controllers spend most of this added time simulating tactical communications. As a result, the fidelity of the BS is seriously degraded, and training effectiveness jeopardized.

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has for some time been addressing the problems associated with SMIs. These problems can be overcome in a variety of ways. Where the interface is between the soldier and hardware, environmental and job designs are being investigated. Where the interface is between the soldier and system's software, artificial intelligence, data management, and natural language communications are being investigated. Where the interface is between the soldier and the hardware and software combined, interactive systems, operational decision aids and speech recognition/synthesis are being investigated.

The focus of the research addressed in this report is speech synthesis/recognition technology and its potential benefits in operational and training systems. This report presents the conceptual design of a computer-assisted simulation of tactical voice communications. This is the first of two research phases. During the second research phase, the system will be built and tested.

REPORT ORGANIZATION

This report is organized into nine sections, of which this is the first. The remaining eight sections will present, in order, the results of this research. These are:

- Voice Technology and Benefits to Army Applications — In this section, the critical role SMIs play in military systems is identified. Using this as a departure point, the potential benefits of voice input/output SMIs to both operational and training systems is addressed. It is pointed out that voice I/O SMIs have yet to be investigated. Therefore, there is a need to conduct research in this area.
- Voice Technology — The purpose of this section is threefold. First, an assessment of the state-of-the-art in voice technology is presented. Second is a survey and evaluation of currently available voice synthesis/recognition and related computer technologies. Third, an optimal approach to developing a computer-assisted simulation of tactical voice communications is determined.
- Definition of Tactical Voice Communications — It was necessary to develop a comprehensive definition of "tactical voice communications." Information and communication theory literature were reviewed and the results documented. Finding no suitable definition in the literature, one was developed. Tactical voice communications are defined in terms of units involved, communication means, communication equipment, and communication nets.
- Taxonomy of Tactical Voice Communications — A prerequisite to the design of a voice I/O SMI is a determination of its voice recognition and synthesis requirements. A taxonomy of tactical voice communications was developed as a means of satisfying this requirement. The approach to its development, the results of a review of both formal Army and research literature on this topic, and a taxonomy itself are reported in this section.
- Computer-Assisted Simulation of Tactical Voice Communications (SIMCOMM) — This section presents the conceptual design of the SIMCOMM. It first addresses SIMCOMM's general requirements in terms of being a standalone system, capable of demonstrating integrated technologies (i.e., voice synthesis, voice recognition and artificial intelligence) and applicable to Army needs (e.g., tactical/small unit, voice-oriented tasks). This is followed by identifying SIMCOMM's application area and operational concept. Functional components of the SIMCOMM are then presented, i.e., soldier requirements (e.g., target representation, voice I/O) and researcher/trainer requirements (e.g., scenario alternatives, data analyses). The basic hardware configuration of the SIMCOMM is then presented, followed by a specification of SIMCOMM's interactive voice protocol requirements. The latter represents a detailed, definitive explanation of SIMCOMM's specific voice synthesis and recognition requirements.

- System's Overview — This section addresses SIMCOMM's operational components/subsystems and specifies SIMCOMM's hardware and software configuration. This section also discusses SIMCOMM's flexibility, modularity, and expansion potential.
- System's Development, Testing and Expansion — The tasks required to carry out Phase II of this research are discussed in this section. Specifically addressed are steps to design, build and test SIMCOMM as well as plans to expand and integrate SIMCOMM with current and future systems.
- Summary — This report concludes with a synopsis and summary of the previous seven sections.

2 — VOICE TECHNOLOGY AND BENEFITS TO ARMY APPLICATIONS

The man/machine interface has been, is, and always will be the key to people reaping the benefits machines have to offer. Early man/machine communications were satisfied by simple means, such as a depressed "on" button or the "clank, clank" of a machine needing repair. These became more "sophisticated" with the introduction of flashing lights, buzzers and bells. With the advent of computing machines, man/machine communication media began to take seemingly bizarre forms, such as crude lamp and tape displays that had to be interpreted by specialists. These were soon replaced by visual-output devices that presented information on pages of hard copy or CRT terminals.

People have become dependent upon machines in every day life. Once, only a small group of people interfaced with machines. Today, most people in developed societies interface with machines on a daily basis. This is especially true in the machine-dependent military environment. It is here that the soldier/machine interface (SMI) is critical. The battlefield of today consists of a vast inventory of machines (e.g., vehicles, radar, communication equipment, weapon systems) with which every soldier must interface at one point or another.

SMI technology has advanced dramatically. These advancements have been necessary given the proliferation of machines in the military, constraints on training resources, the soldiers' learning ability, and the ever-increasing criticality and complexity of modern battlefield machinery. To keep pace with the military's machines, SMI's technological advance must not only continue, but be intensified.

This section investigates the potential benefits of voice technology in Army applications. This topic is best approached by first placing Army applications into three categories: training, operational systems, and research. Having done this, the potential benefits of voice I/O SMIs can be discussed.

TRAINING APPLICATIONS

Approaches to satisfying Army training requirements have changed markedly in relatively recent history. Pure lecture/podium instruction was first supplemented and then nearly supplanted by printed material that emphasized self-paced, individualized instruction. Printed words were then often replaced by illustrations. These soon became outdated with the introduction of audio/visual media to instructional settings. Feeling the full impact of computers, the most recent advancements in training have resulted in CBI (computer-based instruction), CAI (computer-assisted instruction) and CMI (computer-managed instruction)—common, everyday terms within military training communities.

Training and cost effectiveness provided impetus for each of these advancements whether they were what today would be regarded as minor (e.g., lecture to printed material) or major (e.g., CAI, CBI). Training effectiveness, in this context, means achieving learning objectives in the shortest period of time. Cost effectiveness considers many factors including development, operational, and maintenance costs, as well as costs involved in paying the soldiers during training.

The CAI/CBI/CMI systems, applied so often to Army training requirements, are teaching "machines" that have unique SMI needs. It is here that voice technology can best benefit training applications, both in terms of training and cost effectiveness.

Speech is the primary means by which humans communicate. Given this, the potential benefits of voice technology to training effectiveness are considerable for several reasons:

- Training Effectiveness — Using voice I/O SMI, the effectiveness of training systems may be enhanced in terms of comprehension, performance, long term retention, and transferability. Given the natural, fluid, and "real time" attributes of voice I/O SMI, trainees' comprehension (e.g., why, where, when, "big picture") may be enhanced. There is little doubt that the attributes of voice I/O SMIs would increase trainees' performance of voice-oriented skills/tasks, such as those associated with C^3. It may even help non-voice-oriented skills/tasks, especially those that are cognitive in nature. Should trainees' comprehension and performance be enhanced, long-term retention of the skills/knowledges learned could also result. Transferability to actual working environments may also be improved.
- Decreased Training Time — Voice I/O may decrease or eliminate the need for the soldier to learn how to communicate with training systems (e.g., pushing buttons, typing messages or codes on a keyboard). In applications where voice I/O SMIs eliminate the need for the soldier to read written material, an even greater amount of training time may be avoided. Training systems that rely on written materials or a "foreign" communication language (i.e., an artifact of the training system itself), often require trainees to repeat training activities several times before they "get it right." As a matter of fact, current CAI, CBI training systems are designed to facilitate "repeating training activities" (which may indicate a weakness as opposed to a strength in their design). Because voice I/O is a more natural communication mode for most Army trainees, this "repeat phenomenon" may be decreased, and additional training time saved.
- Training Costs — Voice I/O SMIs may decrease training costs not only because of decreased training time requirements. They may also decrease or even eliminate the need for instructors. Other costs savings or avoidances may be realized from the perspective of training development. The expense of developing voice I/O SMIs is relatively low now and decreasing daily. In addition, the cost of voice I/O hardware is also decreasing. (Of course, this is equally true of many other SMI technologies, and assessments of cost savings or avoidances must be tempered accordingly.)

OPERATIONAL SYSTEMS

Voice I/O SMIs could prove tremendously beneficial in operational systems, such as weapon systems, vehicles, and C^3 electronic equipment. For example, voice I/O might reduce skill level requirements of operators. Voice I/O will also reduce the need for hands or vision in soldier/machine interaction. In particular, the systems designer will finally be free to choose what is, from a human factors standpoint, the most appropriate mode of interaction. For example, the designer may now opt for speech synthesis instead of visual displays when:

- The message is simple and uncomplicated.
- The message is short.
- Speed of message transmission is important.
- The message does not need to be referred to later.
- The message deals with events in time or with a particular point in time.
- Visual channels of communication are overloaded.
- The environment is not suitable for the reception of visual messages.
- The user has to move around alot.
- There is a chance the user will be subjected to anoxia or high G-forces.

Other sets of circumstances might cause the designer to select voice input in place of hand- or foot-controlled input devices.

Great strides in operational systems' SMIs have been made in recent years. Indicative of these are the SMIs associated with systems such as Vehicle Integrated Intelligence (VINT²), Information Requirements for Command/Control (IRC²), Vehicle Integrated Defense System (VIDS), and Very Intelligent Surveillance and Target Acquisition (VISTA), to name a few. Voice I/O SMIs are being investigated for each of the SMIs associated with these systems. These investigations indicate the presence of strong beliefs that voice I/O SMIs do have a role to play and can prove beneficial to the Army's operational systems.

RESEARCH

Our discussion of voice I/O SMIs' benefits to training and operational systems has been necessarily tempered with qualifiers such as "may," "potential," and "could." The benefits discussed are logically-based and intuitively feasible. However, the bottom line is "conjecture" because there has been little or no research to determine what, if any benefits could be realized using voice I/O SMIs in the applications addressed. The area of voice I/O SMIs therefore is in desperate need of research.

The hypothesized benefits of such SMIs more than warrant pursuing research in this area. Both basic and applied questions need be addressed—compatability of soldiers and voice I/O SMIs, do voice I/O SMIs increase performance, will voice I/O SMIs increase the effectiveness of training systems while at the same time decrease training costs, can voice I/O SMIs increase the effectiveness/survivability of weapon systems on the battlefield?

Specific questions about the technology itself need to be addressed. For example, it has been known for years what factors detract from intelligibility in analog voice output systems, such as telephones or phonographs. However, it is not known what factors affect intelligibility or listener fatigue in digital synthesizer-based output systems. Many questions about voice input also need to be answered. For example, the previously listed guidelines on when to use audio output have appeared in human factors textbooks for over twenty years. Similar guidelines on when to use speech input have yet to be developed.

Until such questions can be answered, the potential benefits of voice I/O SMIs will never be realized. The research addressed in this report will answer only some of these questions leaving many others unanswered.

SUMMARY

Because speech is the primary means by which humans communicate, why not use it to satisfy SMI requirements? A few years ago, this would have been impossible. However, advancements in semiconductor technology, efficient modeling of the human vocal apparatus, and innovative digital filters based on linear predictive coding (LPC) make the approach feasible and cost effective. Feasibility and cost effectiveness are not the only factors that should cause one to consider voice I/O: Given the Army's dependence upon machines, the complexity and critical role of machines, soldiers' learning abilities, and ever-increasing training demands coupled with diminishing training resources, voice I/O may, in fact, prove to be the <u>best</u> SMI.

3 — VOICE TECHNOLOGY

For many years, the Army has shown great interest in computer-assisted and computer-based training systems. However, none of these systems incorporates automatic speech synthesis and recognition, despite the fact that voice interaction is an ideal medium for many training applications. The Army also recognizes the potential benefits of voice interactive systems to operational systems, e.g., VINT², VIDS, and VISTA discussed previously. Why haven't speech synthesis and recognition systems been used? Very simply, there were no systems good enough for use in these environments. Synthesized speech tended to sound grossly non-human, while speech input systems suffered from poor recognition accuracy. However, these technical problems have been largely overcome. The technology is advancing rapidly, and has reached the point that useful, yet inexpensive, voice input/output soldier/machine interfaces can be built. Indeed, the system being built under this contract is such a system. The purpose of this section is to describe, in layman's terms, how voice interactive systems work.

This section is divided into three major subsections. The first describes the human vocal mechanism, because many snythesis and recognition techniques are based on mathematical models of the human vocal tract. The second subsection describes synthesis, and the third describes recognition. It is important that the reader who is interested solely in recognition also review synthesis, because many speech recognition techniques are based on the vocal tract modeling techniques described in the synthesis section.

HUMAN SPEECH PRODUCTION

The human speech mechanism is complex, and this section presents a comparatively simple model of it. (See Ladefoged, 1975, for a thorough description.) The major physical components of the human speech mechanism include the lungs, the vocal cords, and the vocal cavity, as shown in Figure 3-1.

In the generation of speech sounds, air is forced from the lungs past the vocal cords and through the vocal cavity. The pressure with which the air is exhaled determines the final amplitude, or "loudness," of each speech sound. The action of the vocal cords on the breath stream determines whether the resultant speech sounds will be voiced or unvoiced. The voiced sounds of speech (for example, the "v" sound in the word "voice") are produced by tensing the vocal cords while air is forced from the lungs. The tensed vocal cords interrupt the flow of air, resulting in the release of air in short periodic bursts. The frequency with which these bursts are released imparts pitch to the voice; the greater the frequency, the higher the pitch.

Unvoiced sounds (for example, the final "s" sound in "voice") are produced when air is forced past relaxed vocal cords that do not periodically interrupt the air flow. The sound is generated by audible turbulence in the vocal tract. A simple demonstration of the role of the vocal cords can be had by placing one's fingers lightly on one's larynx, or voice box, while slowly saying the word "voice"; the vocal cords will be felt to vibrate for the "v" sound and for the double vowel (or diphthong) "oi" but not for the final "s" sound.

The sound-generating mechanisms described above produce what is called the excitation signal for speech. There are only three variable parameters (as shown in Figure 3-1) in the excitation signal: its

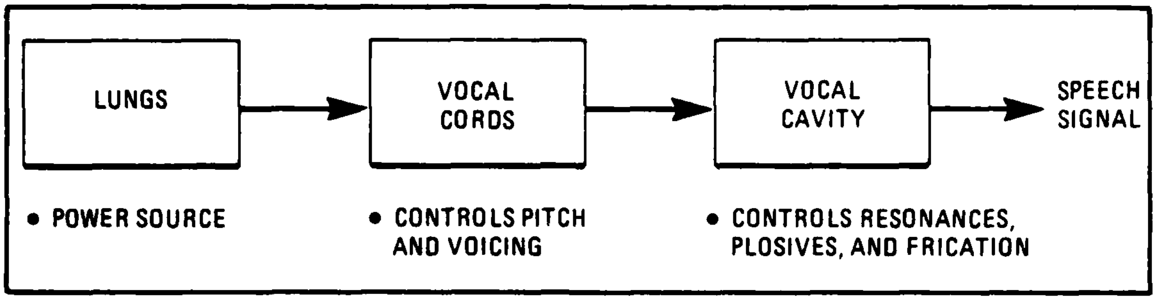


Figure 3-1. **The Major Components of the Human Vocal Mechanism**

amplitude, the proportion of it that is voiced or unvoiced, and, if it is voiced, its fundamental pitch. This can be easily demonstrated. If one were to hold one's mouth wide open, without any movement of the jaw, tongue, and lips, the only remaining changeable characteristics of sound generated by the vocal system would be the above three parameters.

At any given time, the excitation signal will actually contain sounds at many different frequencies. A voiced excitation signal is periodic and the energy in its frequency spectrum lies at multiples of the fundamental pitch, which is equal to the frequency with which the vocal cords are vibrating. An unvoiced excitation signal contains a random mixture of frequencies similar to what is generally called white noise.

The vocal cavity "shapes" the excitation signal into recognizable speech sounds by attenuating certain specific frequencies in the excitation signal while amplifying others. The vocal cavity is able to accomplish this spectral shaping because it reasonates at frequencies that vary depending on the positions of the jaw, tongue, and lips. Frequencies in the excitation signal are suppressed if they are not near a vocal cavity reasonance. However, vocal cavity resonances tend to amplify, or make louder, sounds of the same frequency in the excitation signal. The resulting spectral peaks in the speech sounds are called formants. Typically, only the three or four lowest-frequency formants will be below 5,000 hertz. These are the formants that are most important for intelligibility.

The sounds of human speech can be variously categorized according to the place of articulation, manner of formation, voicing, and the like. For spoken English, a simplified breakdown according to manner of formation would include the following four categories: vowel, nasal, fricative, and plosive sounds.

In the formation of vowels, such as the "e" sound in "speech" and the diphthong "oi" in "voice," the breath stream passes relatively unhindered through the pharynx and the open mouth. In nasal sounds, such as the "m" and "n" in "man," the breath stream passes through the nose. Fricative sounds are produced by forcing air from the lungs through a constriction in the vocal tract so that audible turbulence results. Examples of fricatives include the "s" and "ch" sounds in "speech." Plosive sounds are created when the vocal cavity is completely closed by the lips or tongue and the air pressure built up behind the closure is then suddenly released. The word "talk" contains the plosive sounds "t" and "k." Except when whispering, the vowel and nasal sounds of spoken English are voiced. Fricative and plosive sounds, however, may be voiced (as in "vast" or "den") or unvoiced (as in "fast" or "ten").

Analogies of the human vocal mechanism can be found in those musical instruments that produce sounds by passing an excitation signal through controlled, variable resonators. In the case of trumpets, trombones, and tubas an excitation signal generated by rapid vibrations of the lips rather than the vocal cords can be shaped into different musical notes, or even different sounds, by changing the size and number of the resonating "cavities" through which this excitation signal passes.

Table 3-1

Examples of English Speech Sounds, Classified According to Their Manner of Formation

	Voiced	Unvoiced
Vowel	E as in "Easy" A as in "Answer"	(Not found in spoken English)
Nasal	M as in "man" N as in "Nasal"	(Not found in spoken English)
Fricative	V as in "Vast" Z as in "Zoo"	F as in "Fast" S as in "Sue"
Plosive	B as in "Bat" D as in "Den"	P as in "Pat" T as in "Ten"

SPEECH SYNTHESIS

This subsection reviews speech synthesis techniques. For a more thorough review, as well as an excellent discussion of the human factors issues in voice output systems, see Michaelis & Wiggins (1982).

Figure 3-2 shows a simplified flow diagram of a typical speech synthesizer. This synthesizer has two major components: the source function model that produces the excitation signal, and the model of vocal tract resonant characteristics.

The component that produces the excitation signal has two signal generators. One generates a periodic signal that simulates the sound produced by vibrating human vocal cords. The other produces a random signal that is suitable for modeling unvoiced sounds. Thus, when a synthesizer needs to generate a voiced sound, such as the "e" in "speech," it uses the periodic output from the first signal generator; for the unvoiced "sp" and "ch" sounds in "speech," the synthesizer would instead use the random output from the other signal generator. In some synthesizers, a weighted combination of the random and periodic excitation is used. This can be useful in generating voiced fricative sounds (for example, the "z" sound in the word "zoo"). However, most synthesizers restrict the excitation source so that it is entirely modeled by either the voiced or unvoiced excitation. Alternation of the excitation can be controlled by a two-valued voicing parameter, usually referred to as the voiced/unvoiced decision.

Next, the source function is scaled by an energy or amplitude parameter, which allows the synthesizer to control the loudness. Finally, if the synthesizer is to generate something other than a monotone, it is necessary for the period of the voiced excitation function to be variable. The parameter that controls this is called the pitch parameter. In summary, the excitation signal is specified in the basic model by three parameters: an energy parameter which determines the loudness of the speech; a voiced/unvoiced parameter; and, if voiced, a pitch parameter which specifies the fundamental periodicity of the speech signal.

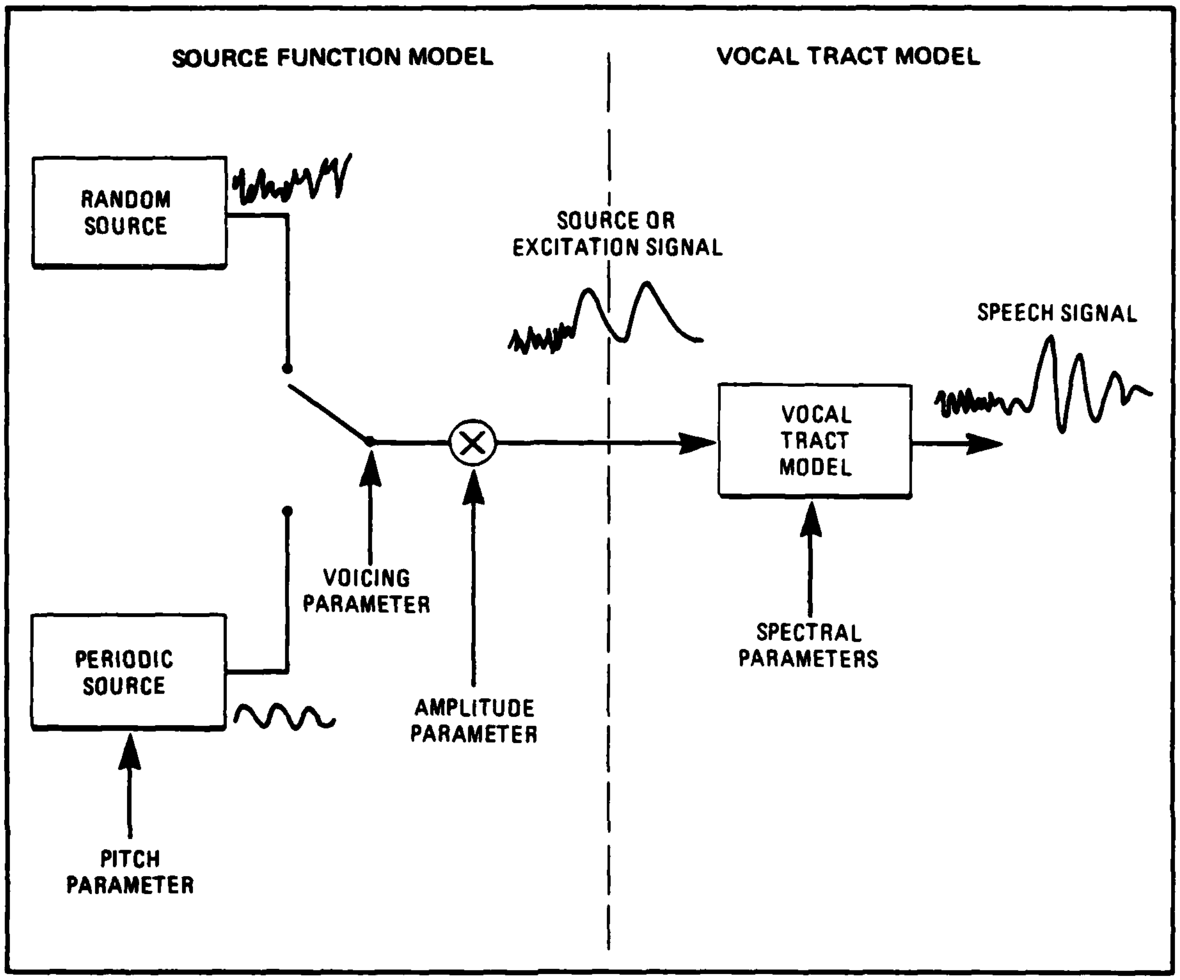


Figure 3-2. **The Basic Speech Synthesis Model Showing the Control Parameters**

The second component of the model in Figure 3-2 is a filtering operation which imposes the proper spectral shape on the artificial excitation signal, just as the human vocal cavity controls the spectral characteristics of its excitation signal. In particular, the resonances of the vocal cavity, which result in spectral peaks in the output speech, can be accurately controlled. These spectral peaks, or formants, are a key spectral characteristic that our hearing systems use to understand speech.

Various techniques have been used to simulate the manner in which the human vocal cavity imposes a particular spectral shape on the excitation signal. One of the first techniques developed uses multiple bandpass filters. The center frequencies of the filters are fixed but an adjustment in the gain of each filter or channel allows the desired spectrum to be approximated. In this manner, a channel synthesizer approximates the vocal tract transfer function by direct spectral measurements (see Figure 3-3). The number of filters required can be reduced if it is also possible to control their center frequencies. By matching the center frequencies to the desired formant frequencies, one can generate synthetic speech with only three or four tunable bandpass filters. Since the controlling parameters in this type of synthesizer are the center frequencies of formants, this is usually called a formant synthesizer (see Figure 3-4).

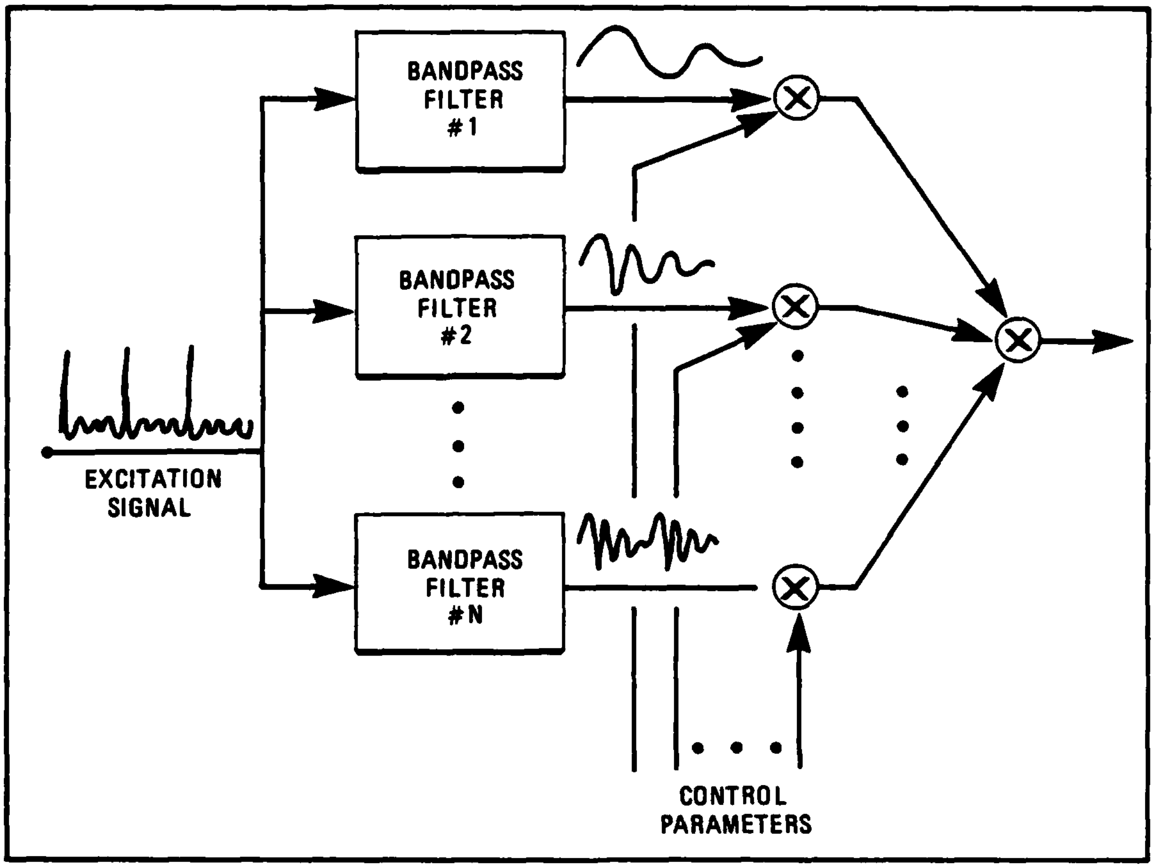


Figure 3-3. **Filter Bank Approach to Vocal Track Model Where the Control Parameters Adjust the Channel Outputs to Appropriate Amplitudes**

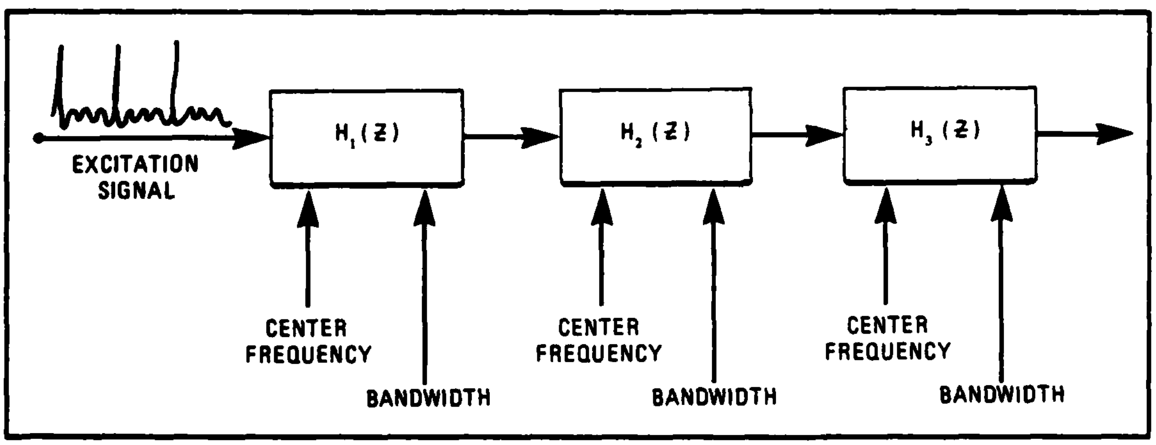


Figure 3-4. **Formant Synthesis Approach Where the Center Frequencies and Bandwidths Control a Cascade of Bandpass Filters**

With the recent advances in digital processing, digital filtering techniques have been widely used to filter the excitation signal. That is exactly what is done in the technique known as linear prediction (see Figure 3-5). In this approach the synthetic speech signal is generated as the output of a filter whose input is the appropriate excitation sequence. Each digital synthetic speech sample can be generated as a weighted linear combination of previous output samples and the present value of the filter input. This yields the following expression for each output sample (S[i]) as a function of previous samples (S[i-1], S[i-2], . . . , (S[i-n]), the prediction weights (A[1], A[2], . . . , A[n]) and the filter input (U[i]):

S[i] =A[i]S[i-1] + A[2]S[i-2] + . . . + A[n]S[i-n] + U[i]

The filter input is the product of the amplitude parameter and the excitation sequence. In linear predictive coding (or LPC), the filter coefficients control the spectral shape of the output signal. As the number of coefficients increases (as n is made larger in the above equation), a larger number of spectral peaks can be approximated. (For a thorough discussion of LPC, see Markel & Gray, 1976. A single-chip LPC-based speech synthesizer is described by Wiggins & Brantingham, 1978.)

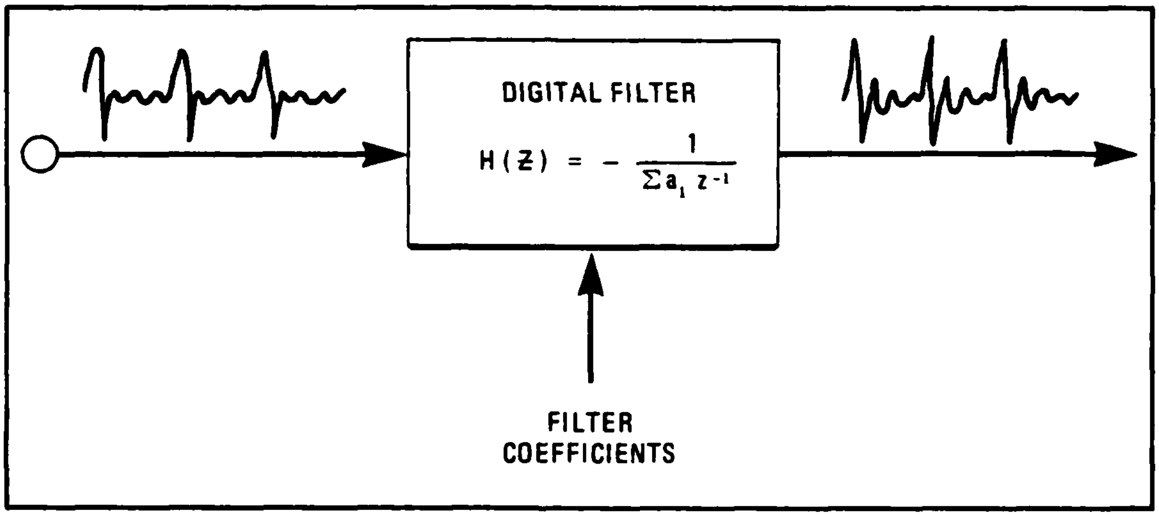


Figure 3-5. **Linear Prediction Uses a Single Higher Order Filter Whose Coefficients are Chosen So That The Transfer Function Matches That of The Vocal Tract**

LPC is useful for recognition as well as synthesis. In TI's recognition system, a spoken word is analyzed to extract, in LPC format, its speech production parameters. These LPC parameters are then compared to existing LPC templates of known words. Very simply, the spoken word is "recognized" if its LPC parameters closely match those of a template. This process is, of course, described much more thoroughly in the section dealing with speech recognition.

Incidentally, the LPC speech encoding model used by TI for both synthesis and recognition is a DoD Standard Algorithm, and is being increasingly used as a data encryption technique for secure tactical voice communication (Tremain, 1982).

Returning to speech synthesis, once the complete set of parameters (amplitude, voicing, pitch, and spectral parameters) has been specified, the speech synthesizer can produce a constant synthetic speech-like sound. Human speech, however, consists of signals with rapidly varying characteristics. It contains many short contiguous segments of voiced and unvoiced speech. Further, the spectral characteristics are constantly changing as the tongue, jaw and lips are moved. For these reasons, the model parameters need to be updated as often as 40 to 50 times each second to generate natural sounding synthetic speech. (Many TI products, such as the Speak & Spell (R), update the speech parameters 40 times per second. However, the simulator currently under construction will update the parameters 50 times per second, thereby producing better quality speech.) In addition, parameter smoothing is usually employed to remove abrupt transitions. This requires that a steady stream of data be supplied to the synthesizer. These parameters can be obtained by an analysis of actual speech, or by an alternative process which is generally called synthesis-by-rule.

ANALYSIS/SYNTHESIS SYSTEMS

Perhaps the simplest and most commonly used method of obtaining the speech parameters is to analyze actual speech signals. In this approach, speech is recorded and a short-time spectral analysis of the signal is performed many times each second to obtain the appropriate spectral parameters as a function of time. A second analysis is then performed to determine appropriate excitation parameters. This process decides if the speech signal is voiced or unvoiced and, when it is voiced, the appropriate pitch values are computed. When the parameters controlling the synthesizer have been carefully determined, the resulting synthetic speech may sound identical to the original. Note that no attempt has been made to reproduce the original time waveform; only the spectral characteristics have been preserved. Figure 3-6 shows a complete analysis/ synthesis system.

In the first step, a person, typically a professional speaker in an anechoic chamber, distinctly pronounces the word or phrase in question. The analog speech signal thus obtained is converted to a digital sequence, generally at a data rate of about 96K bits per second (8,000 twelve-bit samples per second). At this point, of course, the analog signal could be reconstructed with a digital-to-analog converter. In the second step, the digital speech analysis algorithms are used to compute the snythesizer parameters. Note that when these parameters are used in a corresponding speech synthesizer whose output is then converted into an analog signal, a different time waveform is obtained. However, the frequency content of the original speech signal is closely approximated. The speech synthesizer used could be any of several types, including a formant synthesizer or a synthesizer based on linear prediction.

The time trajectories of the speech synthesizer parameters may now be quantized in time and amplitude. Often, the parameters are periodically redefined, typically 40 or 50 times per second, by frames of data that simultaneously update all of the excitation and vocal tract parameters. Generally, coding techniques can reduce the bit rate to the range of 1200 to 2400 bits for each second of speech. (TI's Speak & Spell (R) has a bit rate of 1133 bits per second. However, again in the interests of generating the best possible synthetic speech, the simulator currently under construction will have a rate of 2400 bits per second.) The reduction in bit rate from the original digitized speech may be on the order of 100 to 1. This can represent a substantial savings in memory for voice response applications, or bandwidth for digital voice communications.

TI will be using an LPC based analysis/synthesis system to build the simulator. How well does such a system duplicate human speech? Very well indeed! Figures 3-7 and 3-8 compare original and synthetic waveforms for the words "Synthetic Speech," spoken by a male speaker. Above each waveform is the sound spectrogram. This is a time-frequency-intensity plot of the waveforms and

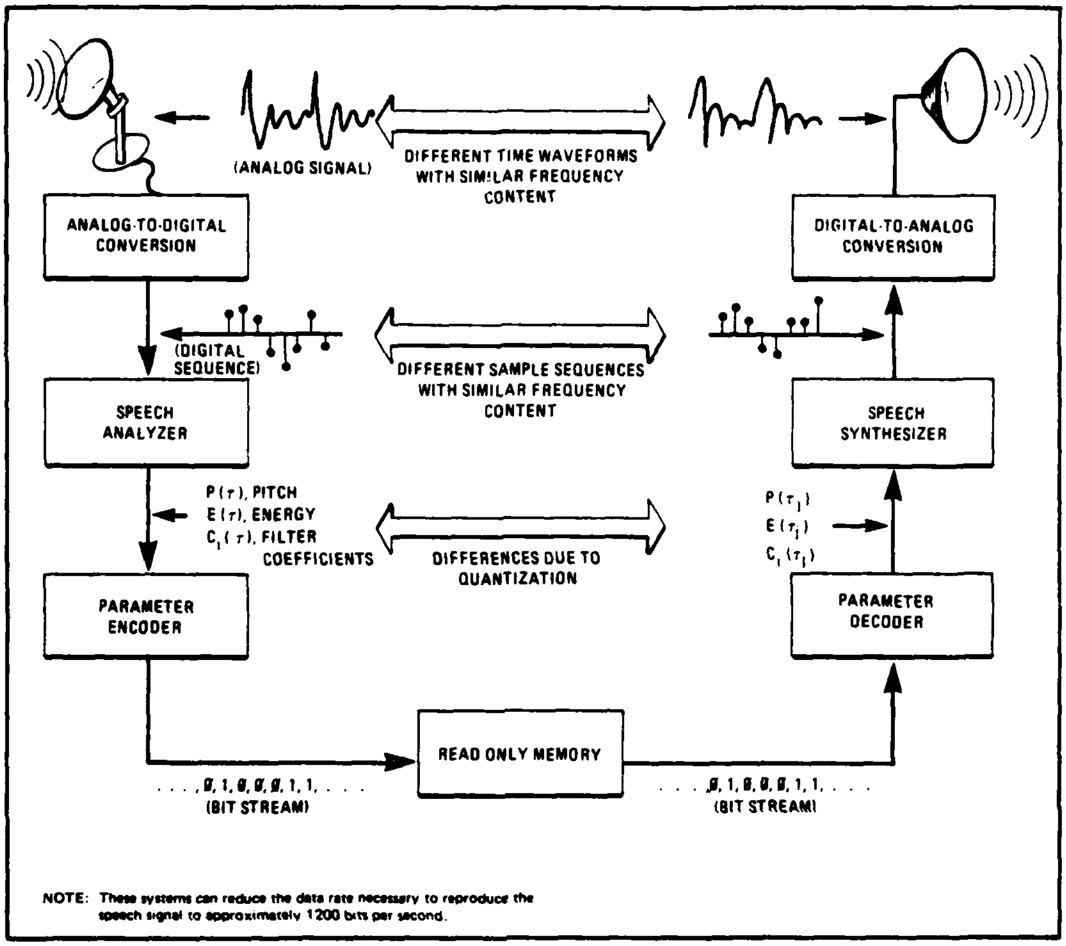


Figure 3-6. **The Digital Analysis/Synthesis System Used for Speech Compression**

clearly shows the preservation of the speech formant structure through the analysis/synthesis procedure. In this example, a spectral analysis of the original speech was performed every 10 milliseconds and the speech was synthesized using a twelfth order LPC model.

From a human factors standpoint, the chief disadvantage of analysis/synthesis systems is that they have a limited vocabulary. They can form messages from only those words and phrases that have been previously encoded and stored in their memories. As a result, the system designer must carefully determine, in advance, all of the words and phrases the system will be called upon to synthesize. (The issue of vocabulary determination is addressed generically by Kelly & Chapanis, 1977; Michaelis, Chapanis, Weeks & Kelly, 1977; Michaelis, 1981; and Section 3.3 of Michaelis & Wiggins, 1982. The issue of vocabulary in tactical voice communications is addressed by this report in Sections 4 and 5.)

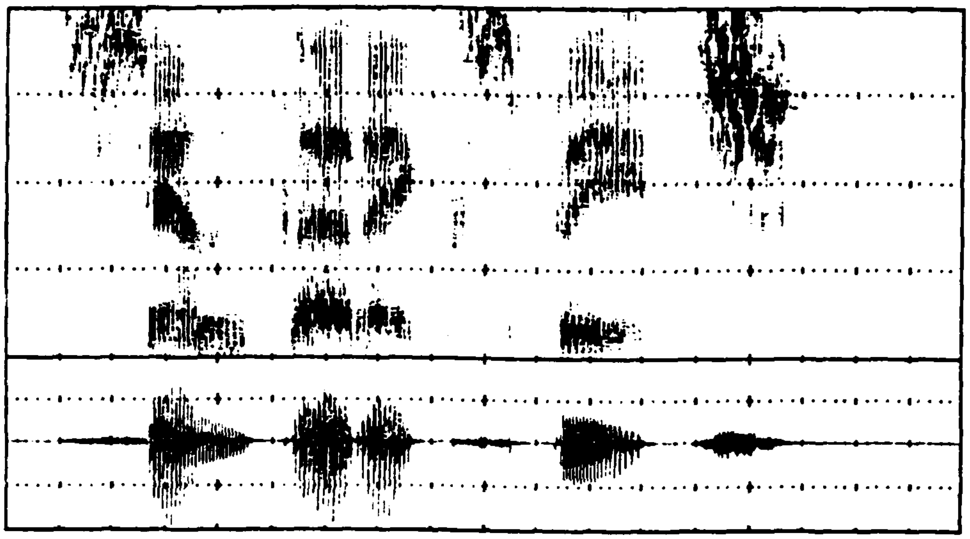


Figure 3-7. **Spectrogram and Waveform of the Original Speech Signal. (The words spoken are "synthetic speech.")**

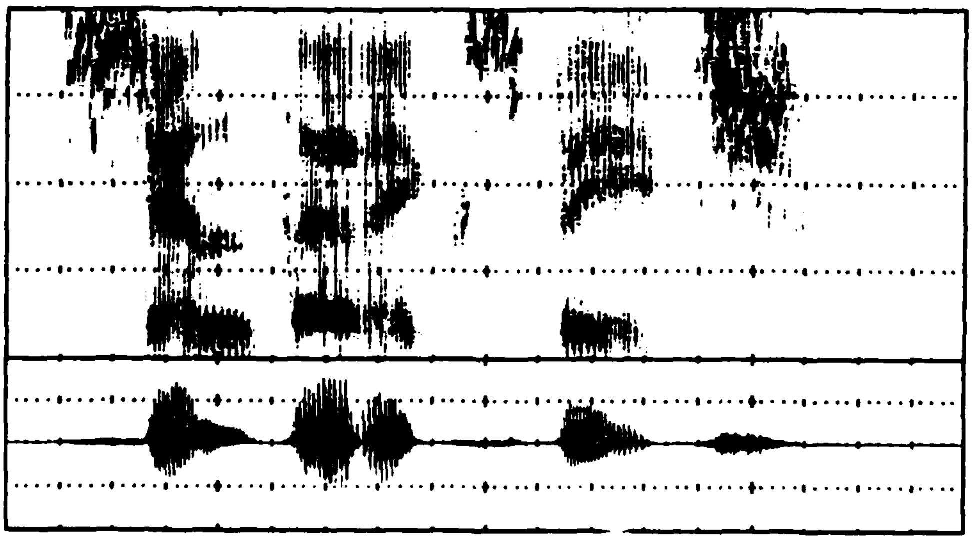


Figure 3-8. **Spectrogram and Waveform of the Synthetic Speech Signal**

SYNTHESIS-BY-RULE SYSTEMS

In cases where vocabulary cannot be determined in advance, a synthesis-by-rule system must be used. Such a system is illustrated in Figure 3-9. When asked to synthesize a particular word, this type of system relies on a set of preprogrammed pronunciation rules to generate the appropriate speech production parameters. The rules vary in complexity from system to system. Some systems simply "sound out" the word using basic letter-to-sound rules. As might be expected, such systems do not always pronounce words accurately. Better pronunciation accuracy is achieved by other, more complex systems that "look up" the pronunciation rules in large internal dictionaries. The top-of-the-line systems also address the problem of coarticulation.

Coarticulation is a modification of pronunciation due to the influence of neighboring sounds. For example, the word "your" in "have it your way" sometimes resembles "chore," as in "have it chore way." In the phrase "sit down," the unvoiced "t" sound is often lost because of the influence of the

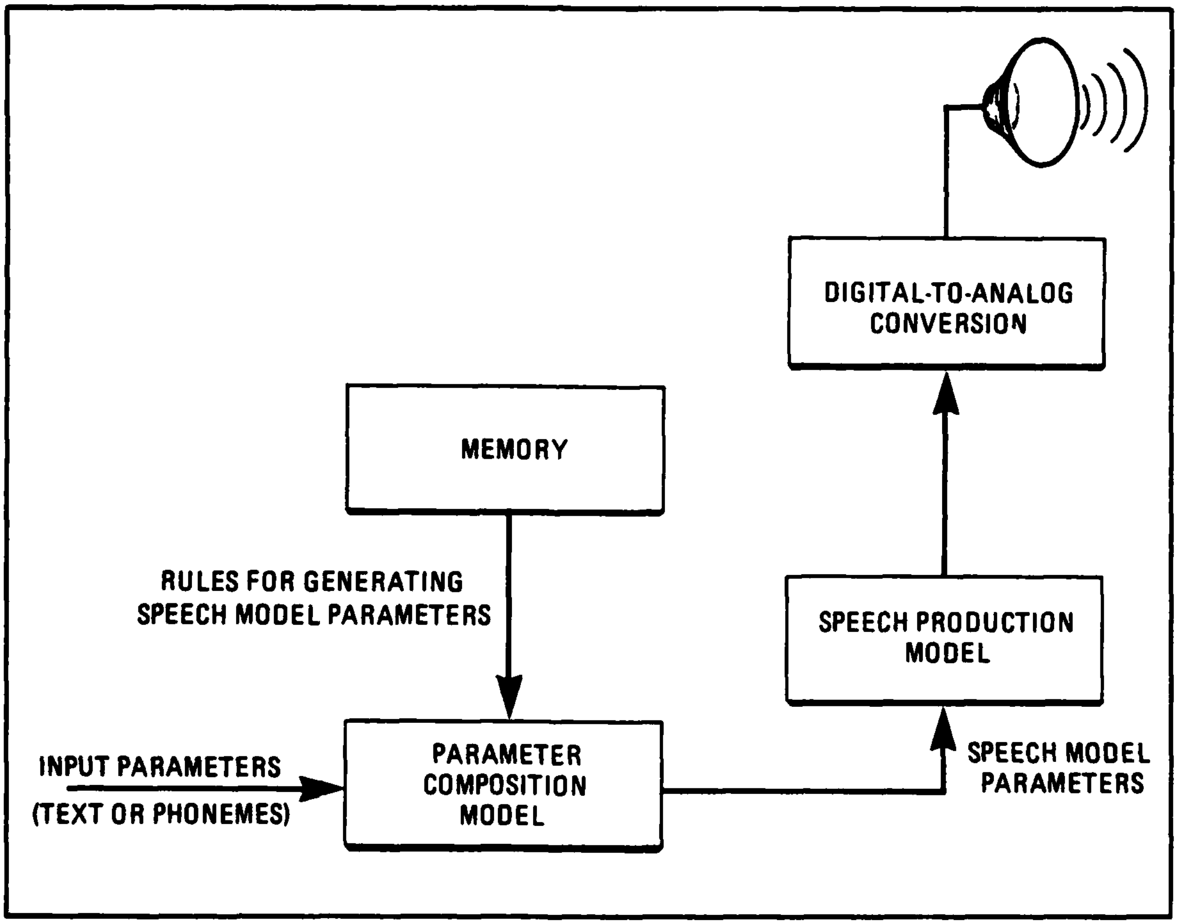


Figure 3-9. **A Synthesis-by-Rule System Generates Speech Model Control Parameters by Applying a Set of Rules to Some Input Parameters, Such as Text or Phonemes**

following voiced "d." In the phrase "used to" the unvoiced "t" causes not only the loss of the preceding voiced "d" but also the unvoicing of the "z" sound in "used." These are examples of a force at work in normal colloquial speech that linguists call "economy of effort"—that is, the tendency to say things the easiest way. Whereas the synthesizer will pronounce "bookcase" as "book" plus "case," in natural speech only one "k" sound is uttered. In instances such as these, the snythesizer, to its detriment, is far more precise than the human tongue.

This problem of coarticulation becomes even more serious when one considers what is generally called "phoneme" synthesis. Phonemes are considered to be the smallest perceptual units of speech, with only 40 to 60 in the English language. However, because of coarticulation problems, natural-sounding synthetic speech cannot be generated by straightforward concatenation of only 40 to 60 different speech segments. In order to sound natural, such synthesizers must modify every elementary speech segment according to a set of coarticulation rules. These modified phonemes are called allophones. Several systems have been developed that synthesize speech by concatenating allophones (see, for example, Lin, Goudie, Frantz, & Brantingham, 1980). Systems of even greater complexity are able to synthesize speech from unrestricted English text (see, for example, Allen, 1976; Kurzweil, 1979). In this process, which utilizes extensive linguistic and syntactic analysis of the text, dictionaries are combined with a set of rules to determine the pronunciation and stress patterns, hence the descriptive name synthesis-by-rule.

The advantages of synthesis-by-rule systems are that larger vocabularies are possible with a fixed amount of memory, and that speech may be generated from more easily interpreted inputs such as text or phonemes. This means that, in some applications, "raw" input data can be supplied by the user at the time of synthesis. An example is a portable phonetic synthesizer that serves as a speech aid for the vocally handicapped (Gagnon, 1978). The disadvantage of synthesis-by-rule systems is that voice quality is generally not nearly as goo' as that produced by analysis/synthesis systems. This is because it is extremely difficult to find appropriate rules for pitch, timing and stress that will produce natural-sounding speech.

SPEECH RECOGNITION

This subsection provides a brief introduction to speech recognition technology. For more thorough discussions, the reader is referred to Doddington & Schalk (1981) and Schalk & McMahon (1982). For a complete discussion of a commonly used speech recognition algorithm, see Itakura (1975).

Before approaching this section, the reader is urged to acquire a firm understanding of human speech production and speech synthesis techniques covered in the preceding subsections.

The system to be delivered will perform speaker-dependent voice recognition. "Speaker-dependent" means that the user must calibrate the system to his own voice before he can use it. The first step in this calibration process is the enrollment mode, in which a speaker-dependent system will prompt the user to speak all of the words it will later be called upon to recognize. As the words are spoken, the system analyzes the speech signals and represents them in terms of time-varying parameters. Several different representation techniques have been used successfully, including direct spectral measurement (mediated either by a bank of bandpass filters or by a discrete Fourier transform), the Cepstrum, and many of the vocal tract modeling techniques described previously.

In systems that use a vocal tract model (such as the system resulting from this research, which will use LPC), the analysis of the speech signal yields the same speech production parameters utilized by synthesizers: voicing, pitch, amplitude, and spectral parameters. Indeed, this portion of the recognition process may be virtually identical to the analysis portion of analysis/synthesis systems, which were described previously and illustrated in Figure 3-6. The speech production parameters thus obtained are used to form templates. These templates will later be used in a pattern matching process to perform speech recognition. However, in many systems, including the system addressed in this report, the templates may first be updated to account for variations in the way the user pronounces the words.

During the update mode, the user is prompted to repeat the words spoken during enrollment. Some systems update the templates by averaging together the original and the new speech production parameters. However, other systems, including the system developed during this research, will first attempt to recognize the words as they are spoken. The recognition process is described later in this section, however, in general, it involves matching the speech production parameters of the spoken word with those stored in the templates. If a correct match is made, the appropriate template is updated and made more robust by averaging in the parameters from the new speech sample. Although this update procedure is not necessary, experiments with TI recognition systems have shown that five updates of the templates reduce the recognition error rate by a factor of three (Schalk & McMahon, 1982). Following this update procedure, the system is ready to recognize speech.

In most systems, including the system resulting from this research, speech recognition is a four-step process: (1) feature extraction, (2) time registration, (3) pattern similarity measurement, and (4) decision strategy. These steps are described below.

- Feature Extraction. During feature extraction, the system analyzes the speech signal in the same manner it did during the enrollment process. The goal is to define the speech signal quantitatively so that it can be compared easily with the signals stored in the templates. Most systems also attempt to reduce the amount of data required to describe the speech signal. They do so by capturing only the important features of the signal, which is why this procedure is often referred to as "feature extraction."
- Time Registration. In general, the incoming speech signal cannot be matched to a template successfully without further manipulations of the signal. This is because people naturally vary the speed with which they speak. This, of course, has a considerable effect on the speech production parameters. Therefore, before trying to match a speech signal to a template, most systems analyze the signal's speech production parameters in order to capture dynamic speech events independent of time. Following that, an attempt may be made to match this sequence of speech events with those stored in the templates.
- Pattern Similarity Measurement. Many techniques have been used to compare the time-corrected speech production parameters with those in the templates. The goal, of course, is to find the closest match. In general, this is done by measuring the similarity between the events in the speech sample and those in the templates. Most systems accomplish this by segmenting the speech signal into frames. (Recall that the speech production parameters for synthesizers are also stored in frames, which are typically 20 to 25 milliseconds in length.) Following that, they do a frame-by-frame comparison of the signal's speech production parameters with those in the templates. It should be noted, however, that the frame length for recognition is typically only 10 to 20 milliseconds; this helps ensure that all important speech events are captured successfully.
- Decision Strategy. The comparison process finds the template whose parameters are most similar to those in the speech sample. However, few systems will decide, without further computation, that this means the speech sample has been recognized. Most systems instead measure the similarity between the two sets of parameters. The amount of similarity must

exceed a threshold value for the speech sample to be considered recognized. In some systems, including the system resulting from this research, this threshold level is adjustable. Such factors as vocabulary size, the type of task to be performed, the system's physical environment, and various user considerations should be taken into account. Setting the threshold level too high causes the system to reject what may often be correct matches, while substitution errors become more common when the threshold is too low. When properly adjusted, the system to be delivered has an error rate of less than 0.5% with a vocabulary of twenty words (Schalk & McMahon, 1982), which compares very favorably with other speech recognition systems (Doddington & Schalk, 1981).

4 — DEFINITION OF TACTICAL VOICE COMMUNICATIONS

There are many possible interpretations of what is meant by tactical voice communications. To prevent any misunderstandings and ensure this report is properly interpreted, a comprehensive definition of tactical voice communications is required.

On the surface, one would think defining tactical voice communications a simple task and certainly one that must have been recorded in print at some point in history. Based upon this assumption, considerable effort was expended in quest of a previously documented definition. Unfortunately, our assumptions proved incorrect and no such definition was found.

To define tactical voice communications, attention was first focused on information and communication theories. These permit us to gain an understanding of the nature of tactical voice communications, but are far from what is required to develop a linguistic database. Therefore, attention was focused on more definitive elements of tactical voice communications, i.e., unit levels, communication means and radio nets. From these more definitive elements as well as information and communication theories the comprehensive definition of tactical voice communications required can be developed. Each of these will be discussed individually.

INFORMATION THEORY

The constant premise of information theory is that the primary purpose of information is to reduce uncertainty present within an organization or system. An organization's or system's demand for information is directly linked to the uncertainty of the environment within which the organization or system exists. This is expressed quantitatively by Ashby's Law of Requisite Variety which states that the communication requirements (amount of information transmitted and received) within a system must be proportionate to the degree of uncertainty or turbulence of its environment if the system is to maintain control. Therefore, systems existing in "friendly" environments characterized by low turbulance, experience little uncertainty and require a minimal amount of information to maintain control. Conversely, systems such as military organizations in combat, find themselves in highly turbulent environments riddled with uncertainties. As such, military organizations require more information to maintain control. Not only is volume or quantity important, but quality (message conveys intended meaning) and effectiveness (elicits the desired impression or response from the recipient) as well as a variety of attributes such as timeliness, accuracy and relevancy.

Most information theorists agree that there are three primary information mechanisms inherent in organizations or systems:

- Coordination by Rules — inherent in systems whose internal conditions and subsequent activities can be predicted and, therefore, preprogrammed allowing the system to operate effectively without communications.
- Coordination by Goals — systems having the ability to specify goals to be achieved by all its participants thereby decreasing the need for coordination (information transmitted or received) employ this mechanism which reduces information requirements.

- Hierarchy — systems unable to develop implicit rules or goals use this information mechanism which increases the systems' information processing capabilities; and as such, information or data are referred to that level in the hierarchy where a global perspective exists for all affected subunits of the system (Galbraith, 1974).

Of the three information mechanisms discussed, the latter is most applicable to tactical voice communications. This is true because of the dynamic nature of the battlefield which renders it impossible for military organizations to predefine conditions and appropriate subsequent activities and therefore employ a coordination by rule mechanism. For the same reasons, military organizations are unable to predefine specific goals to be achieved by all its members which would permit the employment of a coordination by goals mechanism of information processing.

Information theory provides some insights with respect to defining the nature of tactical voice communications. It tells us that information can be characterized by three components (quantity, quality and effectiveness) and a variety of attributes such as timeliness, accuracy and relevancy. In addition, of the three primary mechanisms of information processing, the hierarchy mechanism is most appropriate to military organizations which is the locus of tactical voice communications under study. Though information theory enlightens our understanding of the nature of tactical voice communications, it does little in the way of providing a detailed definition of tactical voice communications.

COMMUNICATION THEORY

Communications has long been recognized of paramount importance to the military. Not only has its importance been recognized but its vulnerability as well. Tactical communications occupy a precarious position subject to degradation (mortality of its communicators, noise, equipment failure), and, as such, there exists ever-present environmental, situational and human variables attempting to sever it (Born, 1981).

Because of the long-recognized importance of tactical communications and the emergence of communication theory, we must attend to its implications. It was originally hoped that communication theory literature would, minimally, provide a definition of communication which would assist in developing a definition of tactical voice communications. The literature review on this topic agreed on few things. However, most of the literature agreed that one dilemma facing communication theory is consensus on a common definition of the term "communication" (Born, 1981). Much of the literature dwelled on this definition dilemma. Lin (1973) identified the most common approaches to defining communications and developed his own:

- Elemental Approach — This approach defines communications in terms of its structural components or elements emphasing communications as a two-way interactive process in that the role of the sender/receiver is a reciprocal relationship as a rule rather than the exception. The elemental approach to defining communications is the most widely recognized.
- Process Approach — In this approach a cognitive perception of communication systems is manifested. Here the effectiveness of a communication system is categorized as being in either a balanced or unbalanced state. The state of the communication system is a function of a person's attitude toward an information source and an issue, and the perceived assertion of the source about the issue. Adopting this approach, communication systems can be expressed in terms of binary values and their "effectiveness" determined.
- Functional Approach — In essence, this approach defines communication in terms of the messages' or communiques' function. Functions, in this context would include informational, instructional and motivational.

- Conceptual Approach — Lin's (1973) own approach to defining communication integrates the previous definitions into a conceptual framework which focuses on the human interaction aspect of communications. The conceptual approach defines communication in terms of four phases. The first of these is the encounter phase involving the linkage between a specific piece of information and the receiver, and the transmission medium. Exchange is the next phase in which the sender and receiver attempt to share and understand the transmitted message. The third phase deals with the influence (both positive and negative) that the communication source may exert on the receiver and is known as the influence phase. Finally, Nan Lin refers to the final phase as the adaption and control phase which prevents the communication system from deteriorating. Unlike the first three phases concerned with undirectional information flow, during this phase feedback is used to establish a two way flow of communication between sender and receiver.

These four definitions of communication were found to be the most prevalent in the literature. Myriad other definitions of communication can be found in the communication theory literature but they serve only to add to the confusion of its study.

Communication theorists' approaches to defining communication appear to fall into two categories, i.e., behavioral and system. Both the process and conceptual approaches to defining communication could be placed into the behavioral category. These behavioral approaches to defining communication can contribute little to either a definition or taxonomy of tactical voice communications. The elemental and functional approaches to defining communication can be placed into the system's category. Identification of the structural components of a communication system emphasized in the elemental approach and function of messages focused upon in the functional approach to defining communication may contribute to the definition of tactical voice communications. Certainly, the structural components of a communication system should be considered in our definition. The function of a communication system's messages should also be considered.

Information and communication theory contribute little to a definition of tactical voice communications. Therefore, it is necessary to develop our own definition for purposes of this research. It has been concluded that tactical voice communications can best be defined in terms of units involved, means of communication, communication equipment and radio nets involved.

UNITS

The focus of this research is tactical communication at and below the company team level. This focus not only assists in defining tactical voice communications, but provides a first-level breakdown or one dimension of the definition required. Specifically, knowing the effort is to focus on company team levels and below permits one to identify the individuals or recipients of tactical communications. These would be limited to seven leader positions found in company teams:

- Armor company team leader
- Infantry company team leader
- Armor platoon leader
- Mechanized/light infantry platoon leader
- Infantry squad leaders
- Mechanized infantry squad leaders
- Tank commanders.

Although each of these leader positions should be present in any company team organization the mix of these positions within a company team will vary depending upon its organization (i.e., tank-heavy, mech-heavy, tank company base balanced team, mech infantry base balanced team). For purposes of defining tactical voice communications, the numbers and mix of these positions is irrelevant. It is only important that we know what they are.

COMMUNICATION MEANS

The leader positions identified as recipients of tactical communications have several means of communication available to them. To further clarify our definition of tactical voice communications, it is necessary to identify what these means are and whether or not all are encompassed in our definition. The Army's FM 11-50, Combat Communications Within the Division (31 Mar 77), as well as other operation FMs (e.g., FMs 71-1, 71-3 and 7-7) identify the following means of communicating available to the positions identified as recipients of tactical communications:

- Radio
- Wire
- Visual (arm and hand signals, smoke, and flags)
- Lights (flashlights, xenon searchlights, flares)
- Panels (used to communicate with aircraft for marking landing zones, drop zones or units positions)
- Sound (whistles, horns, sirens, bells, pyrotechnics, bird calls)
- Messenger

Obviously, our definition of tactical voice communications would not include visual, light, panels or sound. Though messengers as a means of communicating tactical information could certainly be categorized as a "voice communication," it is accomplished on a face-to-face basis which renders it inappropriate for voice synthesis. Therefore, of the seven means of communicating identified in the Army's literature, only two are relevant to our definition—radio and wire.

COMMUNICATION EQUIPMENT

Having identified the individual positions involved and the means of communicating available to these positions, tactical voice communications can be further defined in terms of radio and wire communication equipment available. Identification of the wire and radio equipment involved is important to the fidelity of the tactical voice synthesizer. The voice synthesizer developed will have a simultaneous, two-channel output capability. This will enable the synthesizer to output not only voice communications, but background noises (e.g., explosions, running engines, small arms fire), sounds common to the equipment involved (e.g., breaking squelch, static), and the unique sounds heard over the net when jamming occurs. To achieve this level of fidelity, one must know the characteristics of the equipment involved to determine its sensitivity to picking-up background noises, the idiosyncratic sounds associated with its use and its vulnerability to various forms of electronic warfare (EW) or jamming. Therefore, for purposes of our definition of tactical voice communications, the radio and wire equipment of concern and the individual positions associated with the equipment are identified in the table below:

Table 4-1

Radio and Wire Communication Equipment and Company Team Positions Involved in Tactical Voice Communications

Radio/Wire Equipment	Company Team Position
FM Transmitter AN/PRT-4 and Receiver AN/PRR-9 (Squad Radio)	Used primarily by light and mech infantry squad leaders. To a lesser degree, infantry platoon leaders and company team leaders with infantry squads.
Radio Set AN/GRC-160 linked with APC's Intercom	Mech infantry squad and platoon leaders and mech infantry company team leaders while mounted in APC.
Radio Set AN/PRC-77	Used primarily by mech/light infantry platoon leaders and, to a lesser degree, tank commanders and company team leaders.
Radio Set AN/VRC-64 linked with Tank Intercom	Used by all tank commanders (TCs) regardless of tank type (M1 Abrams, M60A1/A2/A3).
Radio Set AN/VRC-12 linked with Tank Intercom	Armor platoon sergeants, platoon leaders and company team leaders.
Sound-Powered Telephone (TA-1)	Used by all company team positions.

COMMUNICATION NETS

While wire communications are restricted to one channel, the radio equipment involved have multiple channel capabilities. In addition, platoon leaders, in some cases, and company team leaders in most cases, may have more than one radio available to them during combat operations. This is because these positions depend upon multiple channels or nets to conduct operations. They use these nets simultaneously for different purposes. Therefore, our definition of tactical voice communications can be further defined in terms of which of these nets are of concern and, equally important, which are not of concern. To do this, the nets normally used by these positions should be identified first. Field Manual 11-50, Combat Communications Within the Division (31 Mar 77) identifies four different nets which may be used in company team operations:

- Command/Operations Net — This net is used by commanders for tactical control coordination and reporting of tactical data. Orders, coordination, and information of immediate command and operational value are types of traffic commonly passed over this net. Command/operations nets are normally given the highest priority when establishing communication networks in a combat environment.

- Intelligence Net — This net is used as a real-time means for passing intelligence information and spot reports. In addition, the intelligence net is used as the backup for the command/operations net and is given second highest priority for establishment.
- Fire Direction Net — This net links the maneuver elements and supporting indirect fire (mortars and artillery). It is used to pass details of fire missions and adjust indirect fire support. This net is activated as soon as possible.
- Administrative and Logistics Net — This net is used for passing personnel and material supply information and requirements. The priority for establishing this net is less than the previously mentioned nets.

Though all of these nets may be involved in company team operations, the traffic passed over them are not necessarily tactical in nature. Therefore, our definition of tactical voice communications would include only the command/operations and fire direction nets. There is a strong argument for including the intelligence net because it is used as the back-up net for the command/operations net. However, when it is used for this purpose, the communicators could not distinguish between it and the command/operations net in terms of the communications transmitted over it. Therefore, it need not be singled out as a net associated with tactical voice communications.

SUMMARY

Before the research focused on in this research could be pursued, it was of absolute necessity that a precise definition of tactical voice communications be specified. Reviews of both research literature and formal Army documentation provided some, but far from all of the information required to formulate such a definition. For these reasons, it was necessary to develop a definition tailored to the requirements of this research effort.

The working definition of tactical voice communications upon which all subsequent activities related to this research have or will be founded is comprised of six parts as follows:

- Information theory tells us tactical voice communications are characterized by three components (quantity, quality and effectiveness) and a variety of attributes such as timeliness, accuracy and relevancy. In addition, of the three primary mechanisms of information processing, the hierarchy mechanism best describes tactical voice communications.
- Communication theory offers several approaches to defining communication of which two are deemed appropriate for tactical voice communications, i.e., elemental and functional. Identification of the structural components of a communication system emphasized in the elemental approach, and the function of messages focused upon in the functional approach are contributory not only to defining tactical voice communications but development of its taxonomy as well.
- Units involved in the tactical voice communication focused upon in this research enable us to further define the term. Applying communication theory's elemental approach to defining communication, we find the elements involved at the top level in a company team. The elements of the company team can be further broken down into platoons, squads, fire teams and tanks. Elements of the communication system can then be further specified in terms of the positions encompassed unit elements, i.e., company team leader, platoon leaders, fire team leaders and TCs. For purposes of this research, tactical voice communications are restricted to these elements.

- Communication means are restricted to radio and wire. Other communication means (i.e., visual, lights, panels, sound and messenger) are considered outside the scope of what we mean by tactical voice communications.
- Communication equipment, for purposes of the working definition of tactical voice communications, is restricted to the radio and wire equipment specified in Table 4-1.
- Radio nets involved in this research will be restricted to the command/operation and fire direction nets. The other two nets common to company team operations, i.e., intelligence and administrative/logistics nets, are considered outside the scope of this research effort.

Having defined what is meant by tactical voice communications in terms of information and communication theories, military units, communication means and equipment, and radio nets, a working definition of the term results.

5 — TAXONOMY OF TACTICAL VOICE COMMUNICATIONS

Perhaps the most critical task involved in the development of a computer-assisted simulator of tactical voice communications is determining the units or elements by which to classify such communications. In the past, terms such as classifications, categories or constructs have been used to label the products of such efforts. In this research, the term "taxonomy" has been deemed most appropriate.

Before proceeding, it will be advantageous to define what is meant by a "taxonomy of tactical voice communications" and what purpose it will serve. In simplest terms, the taxonomy is nothing more than a construct within which the language of tactical voice communications (as defined in Section 4) can be classified into a set of theoretically relevant categories. Conceptually defined, these categories must also be operationalized in specific terms so they can be used to create a linguistic database of tactical voice communications. To create the linguistic database, source data (such as audio tape recordings of tactical voice communications) must be analyzed and classified in accordance with the taxonomy developed. Accomplishing this, the data can then be subjected to lexical, structural and environmental analyses resulting in the required linguistic database. Therefore, the taxonomy addressed in this section of the report is regarded as the basic building block for the systematic analysis of tactical voice communications which will eventually result in the "dictionary" contained in the voice synthesizer and serve to define the voice recognition requirements of a tactical communications system.

To report this taxonomy, we will discuss the approach applied to its development, the results of a review of both formal Army and research literature on the topic, a definitive taxonomical construct of tactical voice communications and conclude with a summary of this section of the report.

APPROACH TO THE DEVELOPMENT OF A TACTICAL VOICE COMMUNICATIONS TAXONOMY

A three-step approach was employed to develop a tactical voice communication taxonomy. The first and second steps encompassed two literature reviews. Formal Army documentation such as field manuals (FMs) and training circulars (TCs) were reviewed first. The second review focused on the research literature. The usefulness of the results of the literature reviews being limited, the third step involved combining what was useful in the literature and developing a specific taxonomy that would satisfy the requirements of this research.

The strategy most often advocated in the research literature to develop such a taxonomy is to superimpose on the empirical data a theoretical framework defined in terms of a set of conceptual constructs. The abstract constructs then go through several evolutionary stages to define them empirically. A problem then surfaces which Stein and Bruce (1981) described in terms of showing "... a relationship between each operational (empirical) indicator and its theoretical construct" (p. 8). This deductive procedure to the development of a taxonomy, given the scarcity of prior relevant theoretical work, is not feasible.

The development of the tactical voice communications taxonomy reported here followed a somewhat different approach from those advocated by the research literature. In reality, several approaches were used. In this research, we proceeded from the empirical data to the theoretical model and back again. As a result, both an inductive and deductive approach was adopted. This began with a review of formal Army documentation and relevant research literature for purposes of identifying existing taxonomies and/or an examination of the tasks requiring tactical communications. No suitable taxonomies were found. However, in terms of the tasks requiring communications, relevant data were found upon which the taxonomy of tactical voice communications reported here will be based.

REVIEW OF FORMAL ARMY LITERATURE

In accordance with our definition of tactical voice communications (Section 4), the elements of the communication system involved, in terms of units, are those which comprise a company team, i.e., tanks, fire teams, squads, platoons and companies. Elements of the communication system were further detailed in terms of the positions involved (i.e., company team leaders, platoon leaders, squad leaders and tank commanders), radio nets included (i.e., command/operation and fire direction nets), communication means (i.e., radio and wire) and communication equipment. Having defined the communication system in terms of units, positions, communication means/equipment and radio nets, attention was first focused on formal Army documentation in hopes some information would prove beneficial to the development of a tactical voice communication taxonomy.

The Army's FMs and TCs are more or less organized according to unit levels, e.g., companies and squads, and branch or type of unit, e.g., infantry and armor. Therefore, the approach to reviewing formal Army literature centered on the units involved in the communication system under question.

None of the Army literature reviewed included anything resembling a communication taxonomy. Only one of the documents obtained was entirely dedicated to communications. The documents addressing specific levels and types of units had only small sections on communications. These proved most useful to the development of a tactical voice communication definition. However, some clues to the required taxonomy did surface during the review of the formal Army documentation.

As stated previously, the FMs and TCs reviewed are organized by unit level and type. This review will be organized in a similar manner addressing unit levels in ascending order beginning at the squad level.

At the squad level, TC 7-1, "The Mechanized and Light Infantry Squad" (31 Dec 76), has a short section addressing communications. Most of this section is dedicated to describing communication means and how to establish a communication network. The only information in this document relevant to a communication taxonomy is manifested in terms of why the squad leader requires a means of communication — "The squad leader must have a means of communication to control his squad, respond to the platoon leader's instructions, and render reports as necessary" (p. 489). The key words contained in this statement relevant to a communication taxonomy are control, respond to instructions, and render reports.

The next unit level involved is the platoon which is addressed in FM 7-7, "The Mechanized Infantry Platoon and Squad" (30 Sep 77), which states "In order to control your platoon or squad, to report to your commander, to request support, and to respond to orders, you must be able to communicate" (p. D-1). The verbs control, report, request and respond are germane to a communication taxonomy.

At the company team level of operations, FM 71-1, "The Tank and Mechanized Infantry Company Team" (30 Jun 77), states "The team commander must rely on communications to control elements of his command, gather information, distribute intelligence, and coordinate firepower" (p. E-1). The terms in this statement relevant to a communication taxonomy are control, gather, distribute and coordinate.

Because the information contained in the FMs and TCs up to the company team level provided little information relevant to a communications taxonomy, additional formal Army documentation was reviewed. At the brigade level, FM 71-3, "Armored and Mechanized Infantry Brigade Operations" (25 Jul 80), states "Command control is a continuous process in which the commander must find out what's going on; decide what to do about it; follow up to see that it goes well" (p. 7-1). This FM provides additional communication categories in terms of near-real-time information required by brigade commanders, i.e.: "where his subordinate units are; what they are doing; what the enemy is doing as they see it; how the fight is going; what additional support is available . . . ; fuel and ammunition status of units of the brigade; combat vehicle losses of battalion task forces and supporting field artillery" (p. 7-1). Not included with respect to losses or assets were personnel, field fortifications and weapon systems which should have been included.

The only formal Army document that focused entirely on communications was FM 11-50, "Combat Communications Within the Division" (31 Mar 77). There are no counterparts to this document below the division level. This document is less specific than those previously cited in terms of information beneficial to the development of a taxonomy of tactical voice communications. Though less specific, this FM alludes to similar communication factors as the others, i.e., keeping informed, communicating needs and combat achievements.

No specific categories or classifications of tactical voice communications were found in formal Army documentation. However, there was a consensus among these documents' statements relating to the importance and purpose of tactical communications. The terminology used in these statements were also quite similar and provide some insight with respect to what should be considered in a taxonomy of tactical voice communications. This terminology was comprised of the following verbs.

- Controlling/Coordinating
- Reporting
- Responding
- Requesting
- Distributing
- Gathering

The formal Army documentation also alluded to the objects of these verbs which can best be summarized in terms of who (units, individuals, etc.), what (equipment, vehicles, weapon systems, ammunition, etc.), when (time), where (coordinates, topographical feature, phase line, check point, etc.) and why (enemy activity, friendly situation, etc.).

Though a taxonomy, as such, of tactical voice communications was not found in the formal Army literature, an understanding of some factors such a taxonomy should include did emerge from the review of this literature. In addition, with the understanding of tactical voice communications resulting from this review, a more comprehensive and focused review of the research literature was possible.

REVIEW OF RESEARCH LITERATURE

The study of communication does not lack attention from the research community. It has been the subject of research by many disciplines for many years. However, there is remarkably little research documented on the study of tactical voice communications as we have defined it. This conclusion was reached more than three decades ago by Hazell and Leyzorek (1953), again a decade later by Brown (1967) and, more recently by Stein and Bruce (1981) all of whom reviewed research literature on this subject. Most of the research on military communications has focused on the <u>quantification, who, how,</u>

where and when of communications. The what or content of tactical voice communications has been sorely neglected. Unfortunately, it is the content of the communications that is necessary to develop the taxonomy required for this research.

Quantitative analyses of communications are numerous and have involved a number of applications. One early study focused on what was referred to as a study of combat communications (Clarke, Baicker, Cox, Kay, Clement and Bensen, 1953) dealt with the amount of traffic through various telephone links. During the same period, two British studies (Hankin and Love, 1952, and Hankin, 1952) compared transmission frequency and relative cost of manpower for various wire and wireless (radios) communication configurations. Somewhat more germane to the subject of this research was a Johns Hopkins study (Ruark, 1951) focusing on a means of coordinating information involved in joint Naval and Air Force operations. In this study, emphasis was placed on the volume and time associated with tactical communications rather than the content of such traffic.

The first research found to deal with the actual content of communications was reported by Robert Bales (1950) who described a method of categorizing verbal communications. Bales' approach emphasized task and interpersonal behavior resulting in twelve communication categories such as gives opinion, shows antagonism, and shows solidarity. Bales' approach was not applied to a military environment where behavior is dictated by doctrine and tasks of military units which are organizationally unique.

Perhaps the earliest germane study of the content of tactical communications was the work done by Hazell and Leyzorek (1953) who recorded radio and telephone nets of an infantry battalion command post exercise (CPX). Their content analysis of these recordings were reported in terms of three message units, i.e., word, transmission (uninterrupted sequence of words), and conversation (sequence of transmissions). Their analyses showed that the words used were common ones such as I, that, is, you, to, and. Thirty-seven of these words accounted for over half of the content of the communications recorded. Unfortunately, the Hazell and Leyzorek study focused on battalion-level operations, only counted words, did not involve contemporary military TO&E, doctrine or weapon systems.

Another relevant study of the content of communications was conducted by Lennahan (1960) who taped and transcribed command/operations net traffic during three tank battalion exercises. Lennahan identified nine "message classes," i.e., command, weapons/control, intelligence, situation reports, administration/logistics, noncritical, unanswered calls, repeat requests, and obscured. Placing the recorded messages into message classes, Lennahan then quantified the data in terms of number of messages and the individuals involved in the net traffic. The quantification of the net's traffic is irrelevant to the required taxonomy. With respect to his message classes, they closely resemble the taxonomical data drawn from the formal Army documentation and, as such, are no more beneficial to the development of the required taxonomy.

The subject of content analysis of tactical communications was addressed directly by Brown (1967) who studied the content and kinds of information demands on communications necessary within small-unit patrolling operations. While observing both long- and short-range Ranger patrols, observers recorded the time/means of transmissions, message content and sender/receiver of all communications during seven patrols. The data collected on each transmission was then transcribed onto cards which were then placed into two major categories: "commands" referring to the giving and receiving of orders or instructions, and; "information" which pertained to the request for, or giving of information. The messages were then further divided into subcategories resulting in categorization of the messages as shown in Table 5-1. Brown performed several frequency distribution analyses of messages/transmissions, messages/transmissions by sender and receiver as well as modality and techniques. It is the latter which presents problems with respect to the relevancy of Brown's work to this research. The modalities and techniques addressed in Brown's study were visual, sound, lights, messenger, radio and wire. The tactical voice communications focused upon in this research is restricted to radio and telephone. Other problems associated with the relevancy of Brown's work are that it addressed only patrols (similar to squads), was restricted to a single mission (patrolling), and outdated TO&Es, doctrine, and weapon systems. However, Brown's work does provide a model in terms of a structure for and philosophy of tactical communication taxonomies.

Table 5-1

Categorization of Ranger Patrol Communications (Adopted From Brown, 1967)

Commands	Information
Movement (direction, rate, distance, and location)	Movement (direction, rate, distance, and location)
Security (noise discipline, clearing danger areas, cover and concealment)	Security (status and clearance of danger areas)
Fire (weapons and explosives)	Identification (terrain and personnel)
Intelligence (collection, production, and dissemination)	Intelligence (collection and dissemination)
Command and Control (formations, personnel measures, orientation/direction)	Command and Control (status reports, maintenance of orientation/direction, and communication)
Equipment (procurement, placement, utilization, substitution, and preparation)	Equipment (utilization and placement)

During the initial validation of REALTRAIN (a tactical engagement simulation training technique) in Europe, audio tape recordings were made, transcribed and a Communications Index (CI) developed (Root, Epstein, Steinheiser, Hayes, Wood, Sulzen, Burgess, Mirabella, Erwin, and Johnson, 1976). The CI measured the frequency of communications during exercises adjusting the communications frequency based on the number of communicators available per unit of time. The effects of training on the CI was then ascertained. Unfortunately, only the frequency and duration of traffic over the command/operations net were addressed. The contents of the tactical communications were not considered.

Stein and Bruce (1981), in a more recent effort, studied the relationship of communications to combat unit performance. The traffic of the command/operations nets of two opposing units (armored cavalry platoon and an armor platoon) were recorded verbatim during three tactical engagement simulation exercises. A content analysis of these tactical voice communications was then conducted. To perform their content analyses, Stein and Bruce identified eleven "content categories" of the tactical voice communications they had recorded:

- Request orders/guidance (most often, subordinate requests from a superior -"What should I do?","Request permission to move.")
- Provide order/guidance (most often, from superior to subordinate - "Take out that tank!", "You can move your squad now.")
- Request tactical support (requests for and adjustments of indirect fire support -following call for fire procedures)
- Request report/information (subordinate to superior or superior to subordinate - "Give me a SITREP", "When can we expect supplies?")
- Provide report/information (subordinate to superior or superior to subordinate -"The road is blocked!", "You will be resupplied at 0800 hours.")
- Acknowledge message or transmission (common to all communicators - "Roger," "Wilco.")
- Non-tactical (traffic directly related to the control of the exercises - "Where's the controller?")
- Communications facilitation (pertain strictly to net traffic - "How do you copy?", "Lima-Charlie", "Radio Check.")
- Open contact - no content
- Garbled message
- Complex message - multiple content

The Stein and Bruce study is relevant to a taxonomy of tactical voice communications. However, it was derived from a relatively small "N" (three exercises), restricted to specific unit types (air cavalry and armor) and unit level (platoon). For these reasons, it cannot be used "as is" for the required taxonomy of tactical voice communications.

In an effort to develop algorithms for simulating leader behavior in automated battle simulations, a recent study (Hannaman, Underhill, Laurence, and Chambers, 1982) is particularly germane to the development of a tactical voice communications taxonomy. To satisfy the requirements of this study, a general behavior model of the maneuver arms combat leader, irrespective of organizational/unit level or combat arms branch (unit type) had to be developed. The combat leaders focused upon during this effort were company team leaders and below, i.e., the same positions included in our definition of tactical voice communications. Based upon World War II monographs, Vietnam small unit combat action interviews, and a multitude of data (e.g., audio tapes, battle narratives) from tactical engagement simulation exercises (i.e., SCOPES, REALTRAIN and MILES), a General Leader Model (GLM) was developed as shown in Figure 5-1.

Having developed the GLM, Hannaman et al. developed "Leader-Specific Matrices" for all leader positions inherent in a company team. For each leader position, two matrices were developed. The first identified the information received in terms of topic/content, source of the information (specific subordinates & superiors) and the communication modes (radio, wire, face-to-face, written, signals, sound and visual) by which the information may be received. The second matrix specified the actions a leader may take in terms of communications (e.g., requesting/disseminating information) as well as overt actions (e.g., moving, firing a weapon, becoming a casualty). Each of the leader-specific matrices were based upon the GLM.

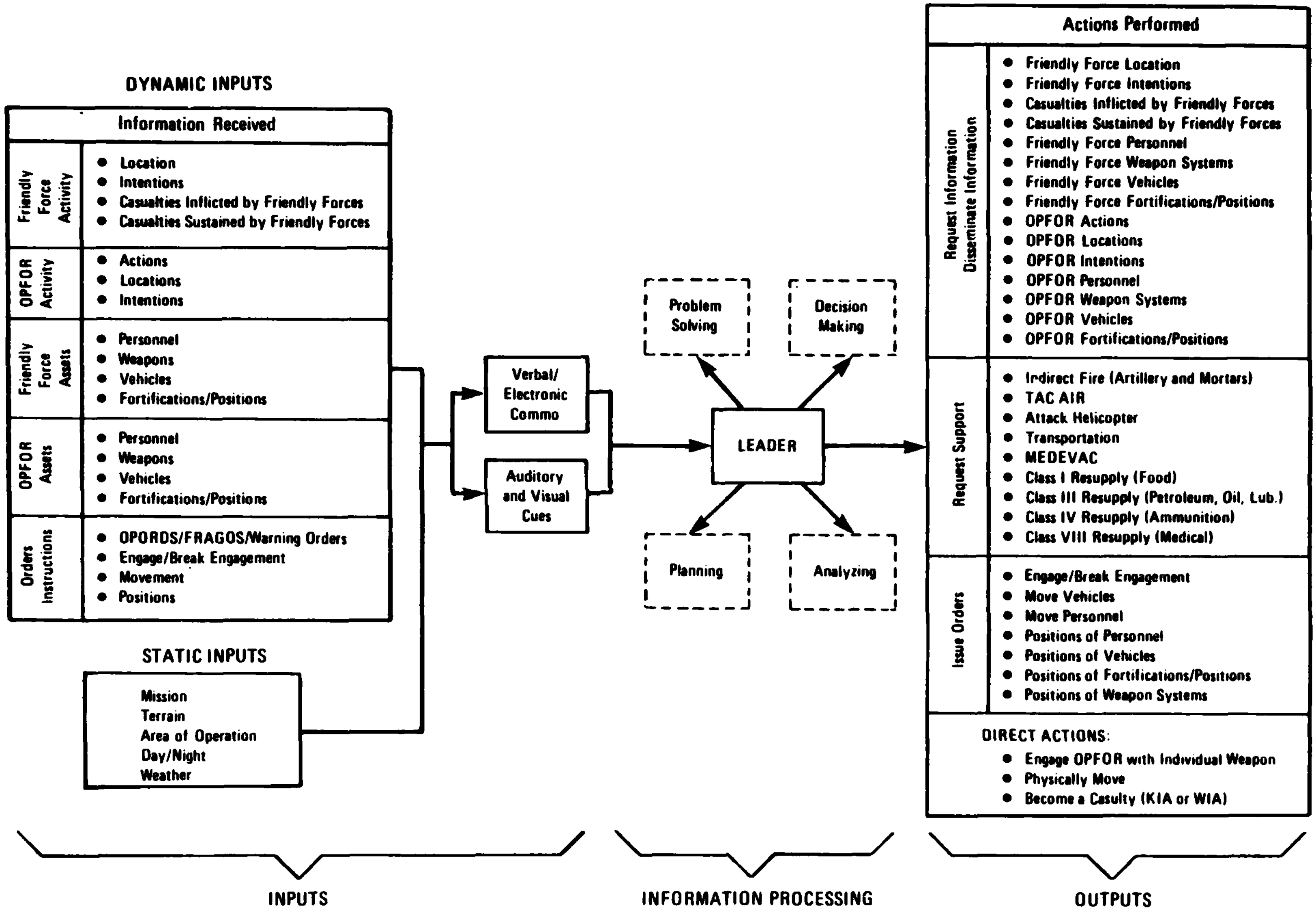


Figure 5-1. **General Leader Model**

To validate the GLM, the leader-specific matrices for light infantry squad leaders were converted to a data collection form for recording all of the information received and actions taken by squad leaders. Twenty-eight squad leaders were then observed while they participated in platoon-level movement-to-contact MILES exercises. To ensure the integrity of the validation, the missions (i.e., point and non-point squads) of the squads whose leaders were observed were varied. The GLM was considered valid in that all information received and actions taken by the squad leaders observed were accounted for in the GLM and leader-specific matrices.

The data collected for purposes of validating the GLM was then analyzed. Several of the analyses performed are of interest to this report. All of the communication means identified earlier were considered in these analyses. Interestingly, the vast majority of inputs (63%) to the squad leaders observed were visual, only 22% involved radio communications and the remaining 15% were comprised of audio (sound) inputs. Most of the communications received by the squad leaders were about friendly force location and orders to move. Most of the squad leaders' own communications dealt with issuing orders to his subordinates to move, cease movement and direction of movement.

The GLM's input and action categories, though not restricted to communications, revolve around tactical voice communications. In addition, it represents a generalizable model of leaders inherent in a company team (regardless of unit level or unit type), is restricted to communications normally associated with command/operations and fire direction nets, and is mission-independent (i.e., applicable to all leaders inherent in a company team regardless of the mission of the team). For these reasons, the GLM proved an asset in developing a tactical voice communications taxonomy.

Many other studies of military communications were reviewed in the course of this effort. These included: the Krumm and Farina (1962) study of B-52 crew communications; Federman's and Siegel's (1965) examination of crew communications in a simulated antisubmarine warfare task which resulted in a 28-category classification of communications which Briggs and Johnson felt literally "defies systematic summary" (p. 33), and; Wood's (1974) history of tactical communications. Regretably, B-52 bombers and antisubmarine warfare have little similarity with company team combat operations.

In conclusion, the current research literature has little to contribute to the development of a tactical voice communication taxonomy. This conclusion is not unique to this paper: in the early fifties, Hazell and Leyzorek (1953) came to the same conclusion; more than a decade later Brown (1967) stated "... most of the studies of military communications have dealt with the structural aspects of message flow, and there has been little analysis of communications content" (p. 5); the conclusions of these researchers were most recently echoed by Stein and Bruce who concluded "The history of communications analysis in a military setting has been laden with efforts to quantify communications ..." (p. 7) and "... little scientific research exists concerning the role, patterning or effect of communications ..." (p.1) within combat arms units.

TAXONOMY OF TACTICAL VOICE COMMUNICATIONS

The taxonomy of tactical voice communications is comprised of seven elements as shown in Figure 5-2, i.e., content classifications, objects, radiotelephone procedural terminology, phonetic alphabet, numbers, interference, and background noises. Each of these elements will be discussed individually.

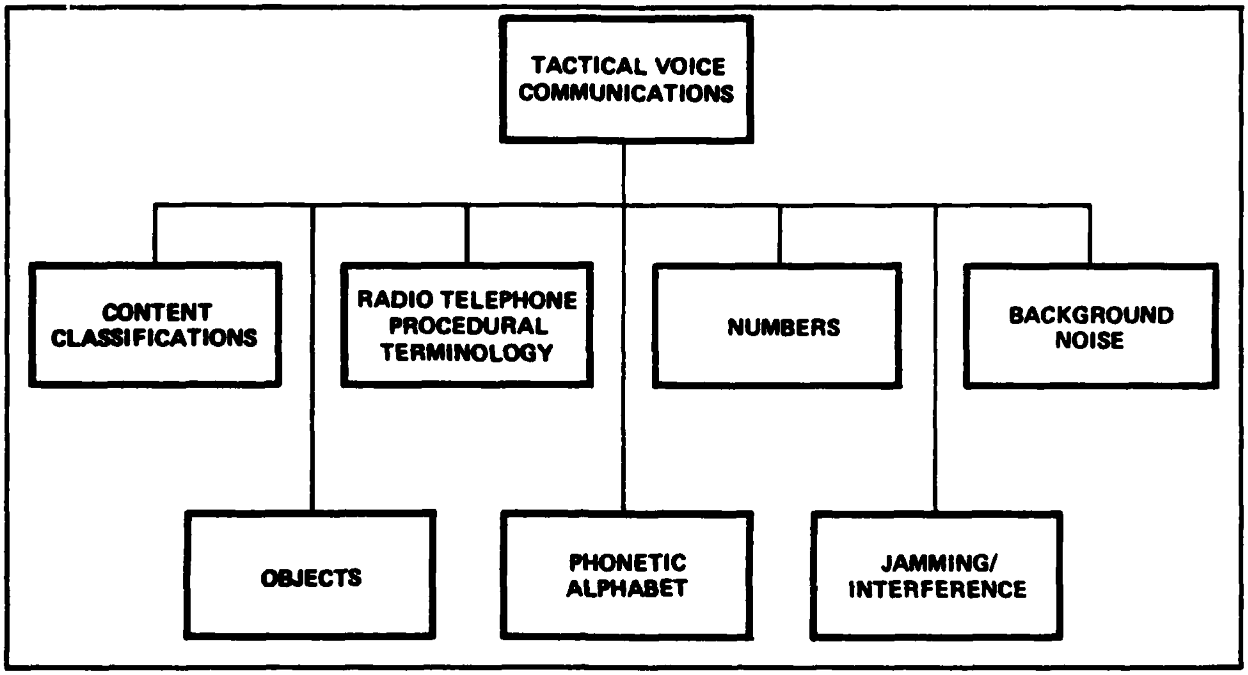


Figure 5-2. Taxonomy of Tactical Voice Communications

Content Classifications of Tactical Voice Communications

Content classifications are defined as the type and subject of messages which might be received by one of the leaders inherent in a company team operation. The message type and message subject categories, based on the General Leader Model illustrated in Figure 5-1 (Hannaman, et al., 1982) as well as the message classification categories developed by Brown (1967) and Stein and Bruce (1981), are listed below in Table 5-2.

As Table 5-2 shows, five message types have been identified. Support Requests and Support Information message types may be regarded as traffic normally associated with the Logistics Net. Because the Log Net was intentionally regarded as outside the parameters of the definition of tactical voice communications presented in Section 4, these message types might easily be regarded as contradictory to the definition. This is a valid assumption until one considers the fact that lower echelon units (i.e., platoons, squads and fire teams) normally do not have access to or use of a Log Net. Therefore, any resupply request messages they may be involved with must be passed over the Command/Operations Net. As this information is passed to higher echelons, it will eventually become part of a Log Net's traffic.

In addition to the message types contained in Table 5-2, twenty-seven message subjects have also been identified. In the case of Information Requests and Dissemination of Information message types, they share fourteen common message subjects. This simply means that the receiver of the tactical

Table 5-2

Content Classification of Tactical Voice Communications in Terms of Message Type and Subject

Message Type	Message Subject
Information Requests and Dissemination of Information	Friendly Force Location Friendly Force Intentions Casualties Inflicted by Friendly Force Casualties Sustained by Friendly Force Friendly Force Personnel Friendly Force Weapon Systems Friendly Force Vehicles Friendly Force Fortifications/Positions OPFOR Actions OPFOR Intentions OPFOR Personnel OPFOR Weapon Systems OPFOR Vehicles OPFOR Fortifications/Positions
Support Requests and Information	Indirect Fire (Artillery and Mortars) TAC AIR Attack Helicopter Transportation MEDEVAC Resupplies [Class I (Food), Class II (Petroleum, Oil, Lubrications), Class III (Ammunition) and Class IV (Medical)]
Orders	OPORDs/FRAGOs/Warning Orders Engage/Break Engagement Movement Positions

Terminology: OPFOR = Opposition Force, TAC AIR =Tactical Air Support, MEDEVAC = Medical Evacuation, OPORDs = Operation Orders, FRAGOs = Fragmentation Orders

communication could either be requested to convey or be provided with information about the subjects identified. Similarly, Support Requests and Support Information message types share nine common message subjects for the same reason.

Objects of Tactical Voice Communication

The next element of the tactical voice communication's taxonomy has been labeled "objects". In this context, objects refers to just that, i.e., the objects which might be included in the message subject categories just discussed. A communication regarding OPFOR or friendly vehicles, weapon systems or personnel may be very general and restricted to these terms, e.g., "Enemy (vehicles, weapon system or personnel) sighted at coordinates 123456." On the other hand, such traffic may be more specific. For example, instead of "enemy vehicle", the communicator may be specific and state "T72 tank," "infantry fire team" instead of "five or six personnel" or "SAGGER missile" instead of "enemy antitank gun." Therefore, this element of the taxonomy would include a vocabulary of specific terms which might be used to describe friendly force as well as OPFOR vehicles, weapon systems and personnel (units). This portion of alinguistic database should be developed based upon the TO&Es (Tables of Organization and Equipment) involved in the scenarios of the application to which the voice synthesizer will be interfaced.

In addition, this element of the taxonomy would include one other object, i.e., terrain features (e.g., stream, hill, depression, road, ridge). These terms are often used in lieu of coordinates when information is either requested or disseminated regarding friendly or OPFOR location, intentions, actions.

Radiotelephone Procedural Terminology

To minimize the probability of the enemy monitoring/jamming a net or "getting a fix" on the location of a communicator, there are two golden rules associated with COMSEC or communication security, i.e., minimize the number and length of communications. At the same time one is minimizing the number and length of net traffic, they must also ensure all of the required information is transmitted.

To satisfy the requirements of COMSEC and, at the same time "get the message across", certain radiotelephone procedural terminologies have evolved in the Army, Air Force, Navy and Marines. These are comprised of commonly used codewords which have distinct meanings and are used to shorten the amount of time involved in tactical voice communications and to avoid any confusion between the transmitter's intended meaning and the receiver's interpreted meaning.

The Communication Appendix of the Army's FM 7-7, The Mechanized Infantry Platoon and Squad (30 Sep 77), as well as other FMs, identifies seven of the most common of these terminologies:

- "OVER" - This term is a codeword meaning "This is the end of my transmission to you and a response is expected. Go ahead and transmit."
- "SAY AGAIN" - This codeword means "I didn't receive or understand your last transmission-repeat it."
- "CORRECTION" - This is used instead of saying "I made a mistake or error in this transmission (or another indicated message). The correct information or version is . . ."
- "I SAY AGAIN" - This codeword means "I am repeating all or a portion of my last transmission."
- "ROGER" - This codeword is used in lieu of saying "I have received and understand your transmission."
- "WILCO" - This codeword mens "I received your transmission, understand it and will comply with (instructions, orders, etc.)."
- "OUT" - This is used when concluding a communication and means "I have concluded my transmission to you and no response is required nor expected."

Minimally, these radiotelephone procedural terms should be included in a tactical voice communication linguistic database. Others, which may be unique to specific scenarios must also be identified and incorporated into the database.

Phonetic Alphabet

Any tactical voice communication linguistic database should include the entire phonetic alphabet so commonly used by the military. This is comprised of a word for each letter of the alphabet beginning with the letter of the alphabet that the word represents, e.g., A = alpha, F = foxtrot, L = lima, P = papa, U = uniform, x = xray and Z = zebra.

The phonetic alphabet serves two purposes. First, it facilitates clarity of net traffic which is sometimes hindered by interference, weak signals and background noises in the vicinity of the communicator. This can be critical when, for example, one wishes to communicate his unit's location using a phase line or check point designated by an alphabetic code. Should the receiver misunderstand and interpret a transmitted "M" as an "N" or "D" as an "E", lives could literally be at stake. The second purpose the phonetic alphabet serves is associated with COMSEC. As stated previously, the golden rule of COMSEC is to minimize the number and length of transmissions—keep them short and simple while at the same time ensuring your message contains all the required information. The alphabetic codes used to designate phase lines and check points is an example of how the phonetic alphabet can be used for this purpose. If phonetic codes weren't used, the only alternative would be to use a six or eight digit grid coordinate. This not only requires a longer period to transmit, it requires precious time to determine what the coordinates are and, if the net is being monitored by the enemy (which one must always assume is the case), one could be divulging their location to the enemy—a deadly situation.

Numbers

A tactical voice communications linguistic database should include all numbers from 0 to 9. Including these, the database will have satisfied all numerical terminology requirements of tactical voice communications.

This is true because military communications rarely, if ever, state the number "318" as "three-hundred-and-eighteen." Instead, they communicate the number by saying "three-one-eight." This procedure is followed to ensure the clarity of the transmission and proper interpretation by the receiver.

Jamming/Interference

The voice synthesizer resulting from this research will have the capability of outputting voice on two channels simultaneously. Of course, radios do not have the ability to receive two, simultaneous communications. Telephones (TA1s), on the otherhand, can receive transmisssions from multiple stations simultaneously. However, the purpose of the voice synthesizer's dual channel output capability is not restricted to the simulation of the idiosyncrancies of the TA1, but also to add fidelity to the simulated radio tactical voice communications. This will be accomplished by the voice synthesizer's ability to output "noises" common to radio communications which have been placed into two categories - jamming and interference.

Enemy forces employ a wide variety of sophisticated radio direction finding (RDF) equipment and knowledgeable, capable communications intelligence (COMINT) analysts who take advantage of our force's loose COMSEC practices. This has become known as electronic warfare (EW), electronic warfare support measures (ESM), electronic countermeasures (ECM) and electronic counter-countermeasures (ECCM). Our potential foes have invested a tremendous amount in each of these which they rightfully regard as a tremendous battlefield resource. These enemy resources are used to disrupt our ability to control our forces on the battlefield. Much of these resources are dedicated to disrupting the communications at the small unit level, i.e., company and below which is the focus of this effort.

The EW, ECM and ECCM activities of enemy forces results in what is referred to as "jamming" our force's communication nets with the command/operation net being one of its prime targets. Jamming manifests itself over the command/operations net in a variety of ways depending on the jamming technique being employed. Specifically, FM 11-50, Combat Communications Within the Division, 31 Mar 77 (p. 3-8) identifies seven ways jamming may manifest itself over our tactical voice communication nets:

- Random Noise - This is synthetic radio noise, random in amplitude and frequency. It is similar to normal background noise and can be used to degrade all types of signals. Operators often mistake it for receiver or atmospheric noise.
- Stepped Tones - These are tones transmitted in increasing and decreasing pitch, and resemble the sound of bagpipes. They are normally used against single channel AM or FM voice circuits.
- Spark - The spark signal is easily produced and is one of the most effective for jamming. Bursts are of short duration and high intensity, repeated at a rapid rate. The time required for receiver circuitry and the human ear to recover after each spark burst makes this signal effective in disrupting all types of radio communications.
- Gulls - The gull signal is generated by a quick rise and slow fall of a variable radio frequency and is similar to the cry of a sea gull. It produces a nuisance effect and is very effective.
- Wobbler - A signal frequency, modulated by a low and slowly varying tone. The result is a howling sound which causes a nuisance effect on voice communications.
- Recorded Sounds - Any audible sound especially of a variable nature, that can be used to distract operators and disrupt communications. Music, screams, applause, whistles, machinery noise, and laughter are examples.
- Preamble Jamming - This jamming occurs when the synchronization tone of speech security equipment is continually broadcast over the operating frequency of secure radio nets. Preamble jamming results in all devices being locked in the receive mode. Preamble jamming is especially effective against radio nets using the current series of speech security devices. Such devices are not currently being employed with radios found at the company team level. However, the rapid advancement of this technology will no doubt result in their employment at this level within a short period of time. For this reason, we have included preamble jamming.

The effects of these jamming techniques, in terms of what might be heard over command/operations nets, must also be included in the tactical voice communications' linguistic database.

In addition to jamming manifestations on the command/operations net, interference noises common to these tactical voice communications must also be included in the linguistic database. These can result from atmospheric conditions, malfunctioning radio equipment, battlefield environment (e.g., nuclear) as well as idiosyncrancies of the radio equipment involved (e.g., breaking squelch).

Background Noises

In addition to jamming and interference being simultaneously output by the voice synthesizer, the fidelity of the voice synthesizer's output can be further enhanced if background noises are also synthesized. These noises can be defined as those present in the immediate vicinity of the transmitter that are of sufficient amplitude and frequency to be picked up by the equipment being used by the communicator. As a result, these background noises are transmitted to the receiver along with the tactical voice communications. These would include, but are not necessarily limited to:

- Small arms fire (e.g., machine guns, M16s)
- Explosions (e.g., impacting artillery, rockets and mortars)
- Talking (e.g., other members of a unit yelling, screaming)
- Engines (e.g., those of APCs, tanks, jeeps)
- Weather (e.g., wind, rain)

These noises combined with those associated with jamming and interference, will add to the "believability", realism or fidelity of the synthesizer's outputs. This can be critical to achieving the purpose or objectives of the simulator to which the synthesizer is eventually interfaced.

SUMMARY

A detailed taxonomy of tactical voice communications has been presented which can be used as the basis for any tactical voice communication linguistic database, e.g., a transcription plan of tactical voice communications regardless of the source data involved. A comprehensive review of both formal Army and research literature has been conducted for the purpose of identifying previous taxonomies of tactical voice communications. These reviews met with only minimal success and a taxonomy of tactical voice communications was developed. The taxonomy has been presented in terms of seven components or elements, i.e., content classifications, objects, radiotelephone procedural terminology, phonetic alphabet, numbers, jamming/interference, and background noise.

The utility of the taxonomy presented is multidimensional. It represents the only known taxonomy of tactical voice communications as we've defined the term. As such, it should prove beneficial to any research or investigation of tactical voice communications in areas such as performance evaluation, training or related voice system requirements. With respect to the research addressed in this report, it will serve as the basic construct within which voice synthesis and voice recognition requirements will be identified and established.

6 — COMPUTER-ASSISTED SIMULATION OF TACTICAL VOICE COMMUNICATIONS (SIMCOMM)

The objective of this research effort is to advance the application of current voice technology to simulation and training systems. It is expected voice technology will enhance these systems' effectiveness and flexibility while simultaneously lowering their operational costs. In this section, we will discuss and describe the Computer-Assisted Simulation of Tactical Voice Communications, or SIMCOMM. It will be described in terms of its general requirements, operational concept and interactive voice protocols necessary to satisfy its voice recognition and synthesis requirements. These discussions will be based upon the definition and taxonomy of tactical voice communications presented in Sections 4 and 5 respectively. SIMCOMM's hardware/software configuration will be discussed in Section 7.

SIMCOMM'S GENERAL REQUIREMENTS

To satisfy the objectives of this research, SIMCOMM must address a set of specific as well as implied requirements. A preliminary set of these requirements was implicit in both the government's request for proposals (RFP) and the TI/HumRRO proposal. This set of requirements matured during the first year of this effort and now can best be described in terms of three areas, i.e., standalone capabilities, technology demonstration capabilities, and application parameters.

Standalone Requirements

A major goal of this effort is to advance the application of voice technology in Army systems. Therefore, it is important that the SIMCOMM not depend upon an interface with any other system. There are several justifications for this standalone requirement. If SIMCOMM were to be designed as an integral part of an existing system, such as Joint Tactical Distribution Information Systems (JTDIS), many time consuming, costly front-end analyses (e.g., structural, lexical, and environmental) would be required. Simply put, the limited resources available renders the feasibility of such an undertaking questionable. As part of a major system, SIMCOMM's benefits/contributions are likely to become masked or buried. As such, it would be difficult and awkward to demonstrate and/or evaluate voice technology.

Another alternative investigated was interfacing SIMCOMM with an Army system currently being developed. This alternative would necessarily link SIMCOMM's developmental milestones with those of the application system being developed. Thus, SIMCOMM's development time might be compressed awkwardly, or delayed unnecessarily. In addition, the drawbacks associated with interfacing SIMCOMM with a fully operational system would also apply to this alternative.

It is necessary to keep in mind that a primary goal of this effort is to educate Army systems' developers about voice technology. Therefore, portability was a major factor strongly influencing the decision to make SIMCOMM standalone: it is much easier to take the technology to the users than bring the users to the technology.

Though SIMCOMM will be standalone in terms of not requiring a "slave" interface with a current or planned system, this will not preclude it being used in that manner. SIMCOMM will be compatible with the most predominant mini/micros as well as large mainframes currently being used by the Army. In addition, it is estimated that no more than 20% of SIMCOMM's available memory will be utilized in its delivered configuration. The remaining memory can be used in a variety of ways including increasing the system's voice synthesis capability with an additional six hours of phrases. (Additional phrases can be added easily by ARI using the standard hardware and software to be provided). If necessary, SIMCOMM's memory can be doubled by adding a second 10 megabyte Winchester hard disc. Alternatively, SIMCOMM could rely on the memory storage devices of any host computer to which it is interfaced.

A final factor influencing the decision to make SIMCOMM a standalone system centered around replication. As delivered, SIMCOMM can be replicated for less than $10K. This is far less than the cost might be if SIMCOMM were a part of a larger system. The low cost of replicating SIMCOMM will increase its potential as a demonstration as well as training system, as well as a demonstrator of future applications.

In summary, five major factors influenced the decision to design SIMCOMM as a standalone system:

- Portability — Making SIMCOMM portable increases the probability of the system influencing and enhancing the application of voice technology in the Army.
- Expansion — SIMCOMM's delivered configuration will not be restrictive. Only 20% of its capacity will have been utilized, permitting considerable expansion. Expansion can be further increased through either the addition of a second Winchester hard disc and/or reliance on the memory storage devices of host computers.
- Interface — SIMCOMM's RS-232 interface will be compatible with most of the Army's mini/micro processors and large mainframes.
- Replication — SIMCOMM's hardware/software costs are minimal allowing it to be replicated at affordable costs.
- Existing/Planned System Constraints — Adverse consequences of its interface with an existing system (e.g., feasibility and/or masking voice technology as a result of being "buried" in a larger system) or planned system (e.g., feasibility and/or compressing/delaying SIMCOMM development) can be avoided.

Considering each of these factors individually as well as collectively leads one to conclude SIMCOMM should be a standalone system if it is to advance the application of voice technology in Army systems.

Demonstrates Integrated Technologies

The system to be developed during this research was to have been restricted to speech synthesis. Though contractually this will remain the case, the SIMCOMM will also demonstrate voice recognition and artificial intelligence. The decision to incorporate these technologies into SIMCOMM was based on the following:

- Human Communication — The bulk of a soldier's communications is by voice. Voice is his most proficient means of communicating. Because of the soldier's natural dependence on speech, non-speech soldier/machine interfaces (e.g., keyboards, CRT's, touch panels) may not always be appropriate or desirable. Therefore, it is desirable for SIMCOMM to provide both speech synthesis and recognition capabilities.

- Artificial Intelligence — Elementary artificial intelligence (AI) techniques will be incorporated into the software. These will enhance SIMCOMM's fidelity.
- Technological Advancements — Voice (synthesis and recognition) and AI technologies have advanced rapidly and dramatically since the initiation of this research. As a result, these technologies can be incorporated into the SIMCOMM with no increase in the research's level of effort (in terms of either manpower or indirect costs).

When these factors were considered, it became apparent that it would be feasible and desirable for SIMCOMM to integrate voice recognition, speech synthesis and artificial intelligence. A sophisticated soldier/machine interface system would result that would not only advance the application of speech synthesis in the Army, but voice recognition and, to a lesser degree, artificial intelligence as well.

SIMCOMM Application Parameters

Having decided that the SIMCOMM should be a standalone system that demonstrates speech synthesis, voice recognition and AI, a specific application had to be agreed upon. To a great degree, the application had been defined in the government's statement of work, but additional details were required before the effort could proceed. Therefore, a set of application parameters were defined as follows:

- Tactical/Small Unit — The government's statement of work made it clear that the application would involve small unit tactical operations. Small unit, in this context, was defined as company team level and below. Tactical simply meant relatively localized, short time frame combat operations against hostile forces. Though these application parameters seem relatively specific, small unit tactical operations are a vast environment encompassing an infinite number of SIMCOMM application possibilities. Therefore, additional parameters were required.
- Voice-Oriented Tasks — Given the original emphasis on speech synthesis and later inclusion of voice recognition, it was clear that SIMCOMM's application should involve voice oriented tasks. These would be tasks or activities in which the soldier is required to receive voice communications (i.e., speech synthesis) and, to a minimal degree, verbally communicate himself (i.e., voice recognition).
- Limited in Scope — SIMCOMM must provide the stimuli for the soldier to communicate as well as cue the system's speech synthesizer. To accomplish this in a standalone system, it is important to limit the intermediate human activities and computer processing. In addition, because the SIMCOMM is to be a "demonstrator," a SIMCOMM session should not be very lengthy. All these requirements dictate that SIMCOMM's application be limited in scope.
- Research Vehicle — The application should provide a good research environment. A primary research area will be human factors and related disciplines. The bulk of studies will center on the soldier/machine interface, for example evaluating the effects of speech input and/or output on performance over a wide range of tasks and environments (e.g., the Very Intelligent Vehicular Information System or VINT[2]). Training research would be another area of interest. These studies might evaluate the effects of voice I/O on training effectiveness and costs (i.e., Costs and Training Effectiveness Analyses or CTEAs).

It should be noted that these application parameters consider both the immediate and future objectives and uses of SIMCOMM. Reflecting these application parameters in SIMCOMM's design will ensure SIMCOMM's short term as well as long term utility to the Army.

SIMCOMM APPLICATION AND OPERATIONAL CONCEPT

SIMCOMM's standalone requirements, need to demonstrate integrated technologies, and application parameters impact both SIMCOMM's application and operational concept. Application is defined as the military topical area or context around which the SIMCOMM will be designed. Operational concept is the means by which this is going to be accomplished in terms of what will be expected of the hardware, software, and human.

SIMCOMM Application

Several candidate applications were identified after extensive review of formal Army literature, such as field manuals and training circulars. The researchers' also drew on their relevant military experience.

The research staff then considered the hardware and software needed for each of these candidate applications. Emphasis was placed on SIMCOMM's standalone and cost requirements. This screening process limited application alternatives dramatically. Of those that remained, "Call for Fire" was determined to be the most appropriate application for SIMCOMM.

"Call for Fire" is the Army procedure for requesting and adjusting indirect artillery or mortar on hostile targets (e.g., troops or "soft targets" and tanks or "hard targets"). In many cases, sole responsibility for this function is assumed by forward observers or FOs from artillery units who are attached to maneuver arms (e.g., armor, infantry units or company teams). However, the vast majority of combat arms NCOs and Officers learn these procedures. This is because there is a limited number of FOs, and their unavoidable mortality rate. Therefore, call for fire is a tactical procedure that should be familiar to the combat arms personnel to whom the SIMCOMM is directed. However, the reasons for selecting call for fire encompassed more than the Army's familiarity with this function. The general requirements discussed previously were also considered:

- <u>Standalone and Interface Requirements</u> — By itself, the call for fire scenario is fairly limited in scope, meaning that it can be implemented on a small, standalone computer system. In addition, several automated battle simulations (BS) currently exist (e.g., MACE) in the Army and many more are planned (e.g., SIMCAT). Each of these battle simulations include the use of indirect fire. Therefore SIMCOMM has the potential of proving beneficial to these systems.
- <u>Application Parameters</u> — Call for fire is inherent in small unit tactical operations. Second, it encompasses a set of voice-oriented tasks. Third, the scope of the communications is limited: it is restricted to two individuals, an established sequence and content of communications, and it involves a fairly restricted vocabulary. Fourth, it is conducive to research in that it lends itself to a variety of objective performance measurements.
- <u>Demonstrates Integrated Technologies</u> — Call for fire represents an ideal domain for demonstrating both speech synthesis and recognition. Artificial intelligence will also be used: as will be discussed later, the actions of the opposition force (OPFOR) will be AI-based.

SIMCOMM Operational Concept

SIMCOMM's operational concept is best described in terms of its two specific groups of users—the subject/trainee and the researcher.

For the subject/trainee, the SIMCOMM application revolves around "Call for Fire." For this, the subject/trainee is given a topographical map and a set of Communication Electronic Operating

Instructions (CEOI). He is asked to detect targets and engage them with indirect fire. The CEOI will be identical, in format and content, to what would be provided in a real tactical environment. As targets are presented, SIMCOMM will display the target itself, impacting fire, cause targets to react in an intelligent manner, and, of course, permit the subject/trainee to call for fire. In addition, SIMCOMM will promote proper radio COMSEC by enforcing authentication procedures and simulating radio-frequency jamming. From the subjects'/trainees' perspective, SIMCOMM's principle functions are:

- Target Representation — The system will present targets in such a manner that it is both feasible and appropriate to engage them with indirect fire. The targets will be a combination of vehicles such as tanks and APCs as well as personnel. In most angles of view and distance, the targets will be shown as silhouettes. This is not expected to decrease fidelity or hinder the subject's/trainee's ability to identify targets.
- Voice I/O — SIMCOMM will include the TIPC speech peripheral which performs both speech synthesis and recognition. The vocabulary to be recognized, as well as spoken by SIMCOMM, reflect formal Army procedures associated with call for fire. SIMCOMM will not require the subject/trainee to type on a keyboard or use any other mechanisms normally associated with interfacing with a computer.
- Impacting Indirect Fire — Once SIMCOMM has provided the stimulus(i) and the subject/trainee has called for fire, the system will present the results of the request for indirect fire by displaying the indirect fire impacting where the subject/trainee called for the fire (i.e., at either the coordinates given or location of the adjustment provided by the subject/trainee). This of course must be represented in relationship with the target location which may have remained stationary, continued to move or stopped and moved several times between the time the fire was requested and actually impacted.
- Intelligent OPFOR Reaction — Once the subjects'/trainees' initial fire request has impacted, the opposition force (OPFOR) must react in a realistic manner. This "intelligent OPFOR reaction" requirement necessitates that SIMCOMM provide the ability for the OPFOR to sense they are being engaged and react accordingly. This means the simulated OPFOR must first sense the impacting fire as well as the topographical features (e.g., woods, depressions, hills) in its immediate vicinity. The simulated OPFOR would then move rapidly and directly to the closest location providing cover or concealment.
- Subsequent Activities — After the subject/trainee has detected the OPFOR, engaged with indirect fire, and the OPFOR has reacted, SIMCOMM will permit the subject/trainee to follow-up their initial actions. SIMCOMM will permit subsequent adjustments to initial indirect fire requests until the target(s) are either destroyed or immobilized or the subject/trainee is destroyed (in the simulation).
- Communication Security (COMSEC) — SIMCOMM will encourage and promote good COMSEC. SIMCOMM will initiate authentication procedures when appropriate. In addition, if the radio communication from the subject/trainees is too long, SIMCOMM will simulate various types of jamming. The subject/trainee must know and use authentication responses, recognize jamming stimuli, and choose alternate frequencies necessary to continue communications. Authentication responses and alternate frequencies will be provided to the subject/trainee in the form of a CEOI as mentioned previously.

For the researcher, a set of different functional requirements is needed. To facilitate researchers' needs, SIMCOMM's functional requirements include:

- Scenario Alternatives — The researcher will be able to modify the basic call for fire scenario. These modifications, or scenario alternatives, will include the number, type and rate of movement of the targets. Thus a variety of scenarios can be developed from a relatively small set of OPFOR variables.

- Linguistic Database Modifications — SIMCOMM will permit the researcher to modify its voice inputs and outputs, including the addition and deletion of words and phrases.
- Data Analyses — The researcher will be permitted to specify many types of data the system should store, such as the number of requests, rounds expended, and response time. It should be noted that SIMCOMM is configured around a Texas Instruments Professional Computer (TIPC), which can also be used for other purposes. For example, the TIPC may be programmed by the researcher in any of several different languages, and may even be used to run statistical packages.
- Hardcopy Output — To support researchers' needs, the SIMCOMM will have the ability to provide hardcopy output of things such as research data, programs and performance data.

Hardware Configuration

To satisfy SIMCOMM's users' functional requirements, specific hardware configurations are dictated. The two primary factors considered in determining what hardware should be used were fidelity and function. Fidelity factors related only to the hardware configuration associated with subject/trainees. Functional factors, of course, were a prime consideration for both researchers and subject/trainees.

The subject/trainee SIMCOMM hardware configuration is shown in Figure 6-1. There are two major hardware components encompassed in the subject/trainee SIMCOMM configuration. The first

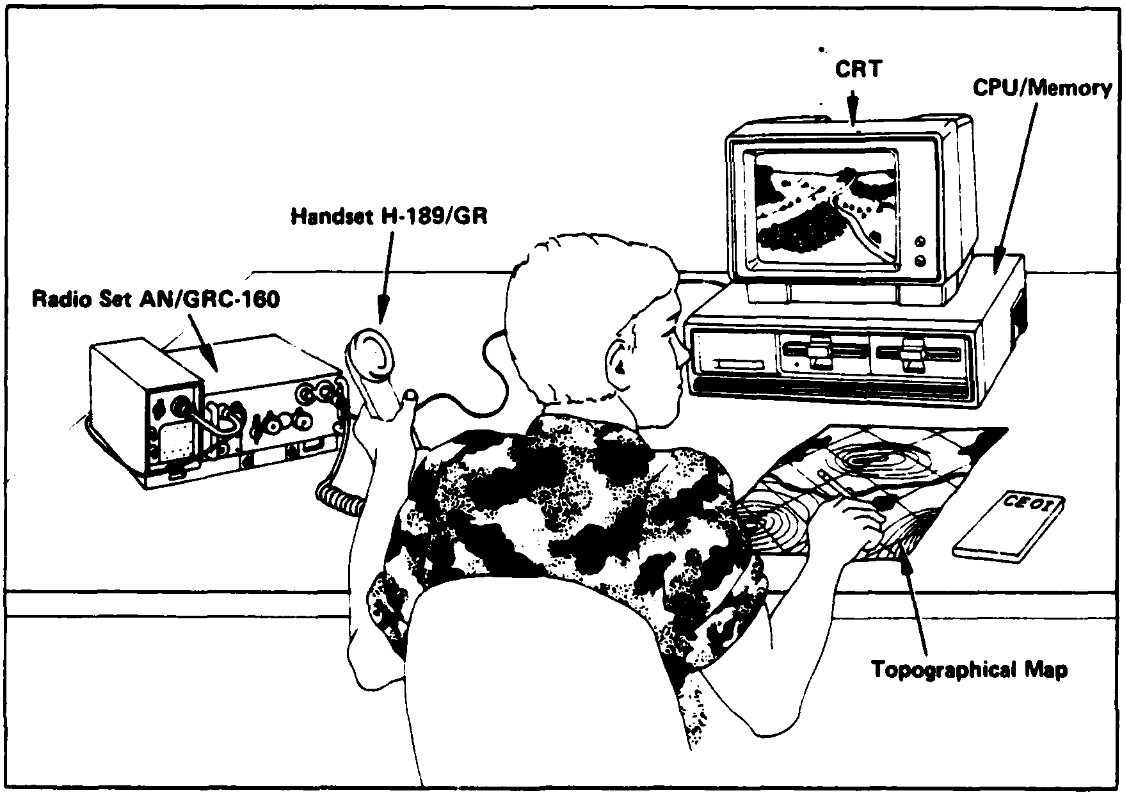


Figure 6-1. SIMCOMM Configured for Subjects/Trainees (Note absence of keyboard)

of these is the Texas Instruments Professional Computer. Its CRT will present a visual representation of the terrain and target(s) to the subject/trainee. The terrain will be represented realistically by presenting what a call for fire requester would see from a strategically located observation post. This visual representation will be identical to the area covered by the topographical map provided to the subject/trainee. Targets (i.e., vehicles, personnel and weapon signatures) will also be presented on the CRT—overlaid on the terrain representation. These targets will present themselves as both stationary and moving. In addition, impacting artillery will also be displayed on the CRT.

The second major hardware component of SIMCOMM's subject/trainee configuration is a radio set. This is shown in Figure 6-1 as an AN/GRC-160 with a H-189/GR handset. This radio and handset will be nothing but an empty shell. Its only function is to permit the subject/trainee to verbally communicate with SIMCOMM. To add perceived fidelity to the system (from the perspective of the subject/trainee), it was decided to use "real" radio equipment. It should be noted that although an AN/GRC-160 is depicted in the figure, another radio might be used, e.g., an AN/PRC-77 with H-189/GR handset.

SIMCOMM's researcher hardware configuration is illustrated in Figure 6-2. Before describing its components it is important to note that the TIPC illustrated is the same one that will be used in the

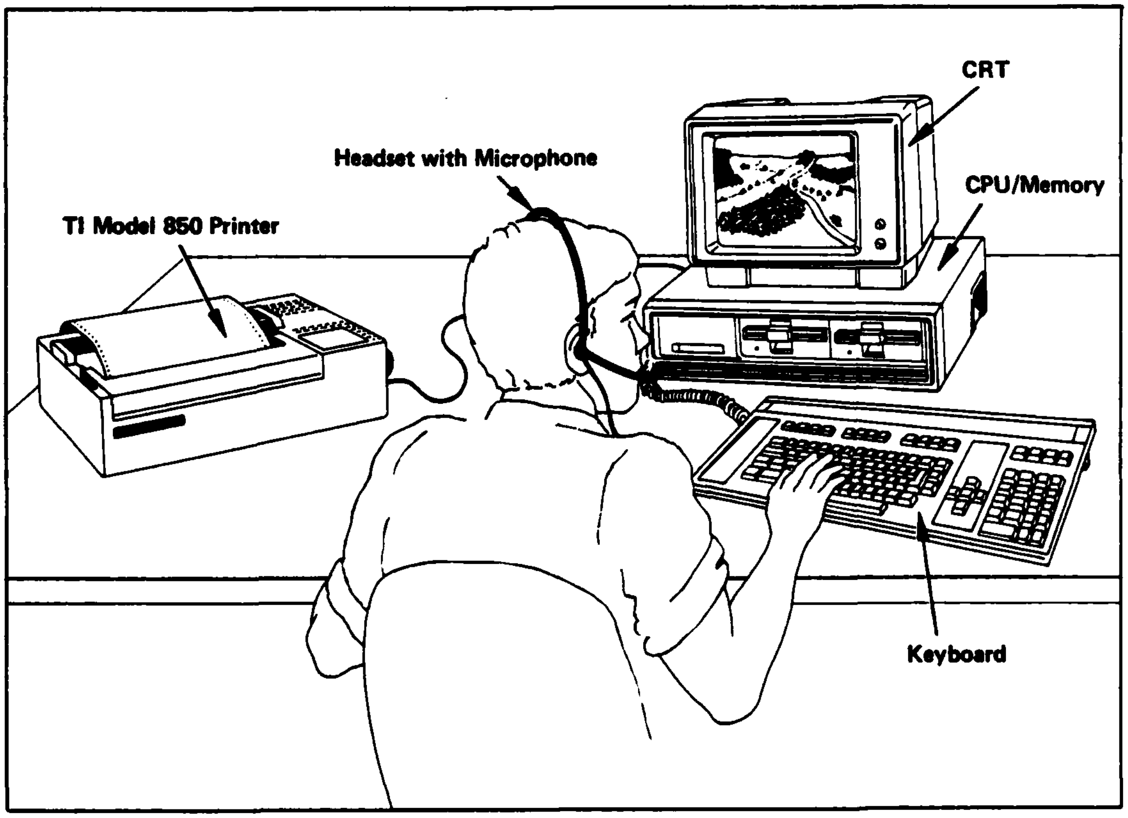


Figure 6-2. SIMCOMM Configured for Researchers/System Developers

subject/trainee configuration. In the researchers' configuration, the radio set has been disconnected and a keyboard, headset with microphone and printer added. The keyboard will permit the researcher to add, delete and modify SIMCOMM's software. The keyboard's layout is shown in Figure 6-3. The headset with microphone will permit the researcher to add, delete and modify SIMCOMM's linguistic database (i.e., voice synthesizer). The printer will provide hardcopy outputs the researcher may desire.

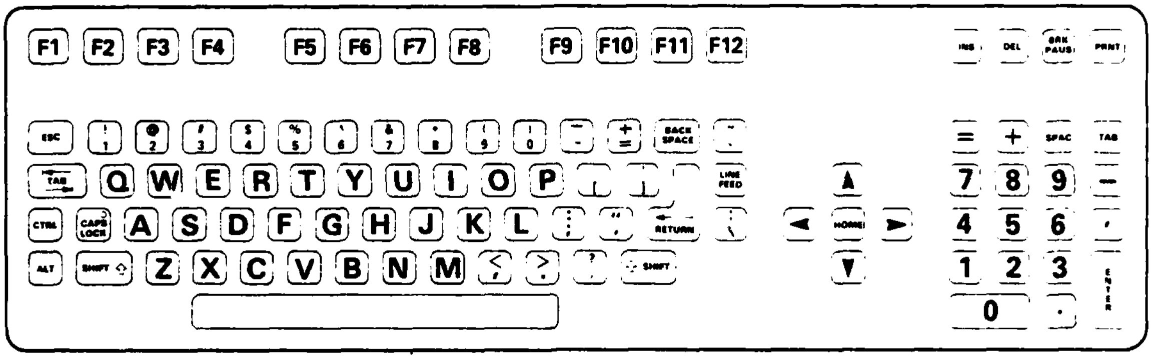


Figure 6-3. SIMCOMM Keyboard

The Texas Instruments Professional Computer will consist of a basic TIPC, with a 320K floppy disk, 256K memory, and monochrome monitor. The options provided will include a 10 Megabyte Winchester disk, 8-color/gray scale graphics board, RS-232C interface board, Speech Command board, and a TI Omni-850 printer. The Speech Command board will permit SIMCOMM to synthesize and recognize speech. The Omni-850 printer will provide for hardcopy of researchers' data, programs, and so on. Aside from the printer, this configuration consists of three units: a keyboard (to be used primarily by the researcher), a monochrome monitor, and the chassis containing the CPU, memory, disk drives, and the various boards to provide speech, graphics, and RS-232 interfaces.

SIMCOMM'S INTERACTIVE VOICE PROTOCOLS

SIMCOMM will incorporate three basic scenarios, i.e., call for fire, communication security (COMSEC), and jamming. Given this understanding, protocols based upon available formal Army documentation (e.g., Field Manuals and Training Circulars), have been developed encompassing these three areas. To satisfy the requirements of the three protocol areas, a total of seven protocol modules (which are referred to in this paper as either "requests" or "sequences") were developed:

- Communication Check (Commo Check) Sequence
- Initial Adjust Fire Request
- Subsequent Adjust Fire Request
- Fire for Effect Request

- Shot/Splash Sequence
- Authentication Sequence
- Jamming Sequence

Each of these and their probable sequencing are illustrated in the flowchart in Figure 6-4. The Authentication Sequence satisfies SIMCOMM's COMSEC requirements. The Jamming Sequence satisfies SIMCOMM's jamming requirements.

These modules are individually explained in detail in the following subsections. These descriptions are presented using a combination of narrative, flowcharts and tables. Each narrative presents a general description of the SIMCOMM module. The flowcharts present the basic logic associated with each module. A table for each module is then presented which identifies the specific synthesis and recognition requirements, their relationship and sequencing for each module. In conclusion, a set of general and specific rules that might be associated with the SIMCOMM modules as they are defined is presented.

The contents of this section should be viewed as a conceptual design for the SIMCOMM. From it, the feasibility of both the voice synthesis and recognition requirements has been determined.

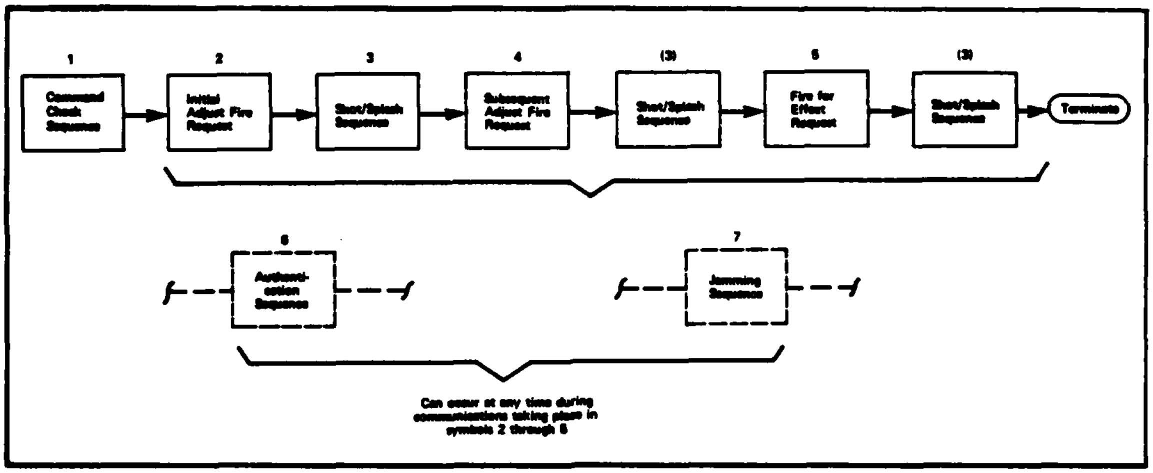


Figure 6-4. Generic Call for Fire Communication Protocols and Their Probable Sequencing

Initial "Adjust Fire" Request

The initial fire request will always be initiated by the system's user. There are several manners in which he could do this, i.e., begin with a fire for effect, adjust fire (e.g., from TRP, from last mission). System rules will dictate that the initial mission be an "Adjust Fire" mission and the target location be given using a six-digit grid coordinate, a direction (in mils) and distance (in meters of 50-meter increments).

Figure 6-5 presents the sequence of communications and basic logic for the initial adjust fire request. Each symbol in this flowchart requiring explanation has been numbered. Each numbered box will be discussed individually with voice recognition and synthesis requirements identified in Table 6-1 immediately following the flowchart.

Symbols one through six in Figure 6-5 correspond to the six major components of call for fire request, i.e., Identification of Observer, Warning Order, Target Location, Target Description, Method of Engagement, and Method of Control.

It should be noted that grid coordinates, distances, directions and call signs can be minimized through careful system design. For example, the topographical map used and the location of targets presented on the screen can be designed in such a way that the first, second, fourth and fifth digit of a six-digit coordinate are always the same, and call signs restricted to two words such as "Bravo six" and "Yankee two." This would minimize recognition requirements.

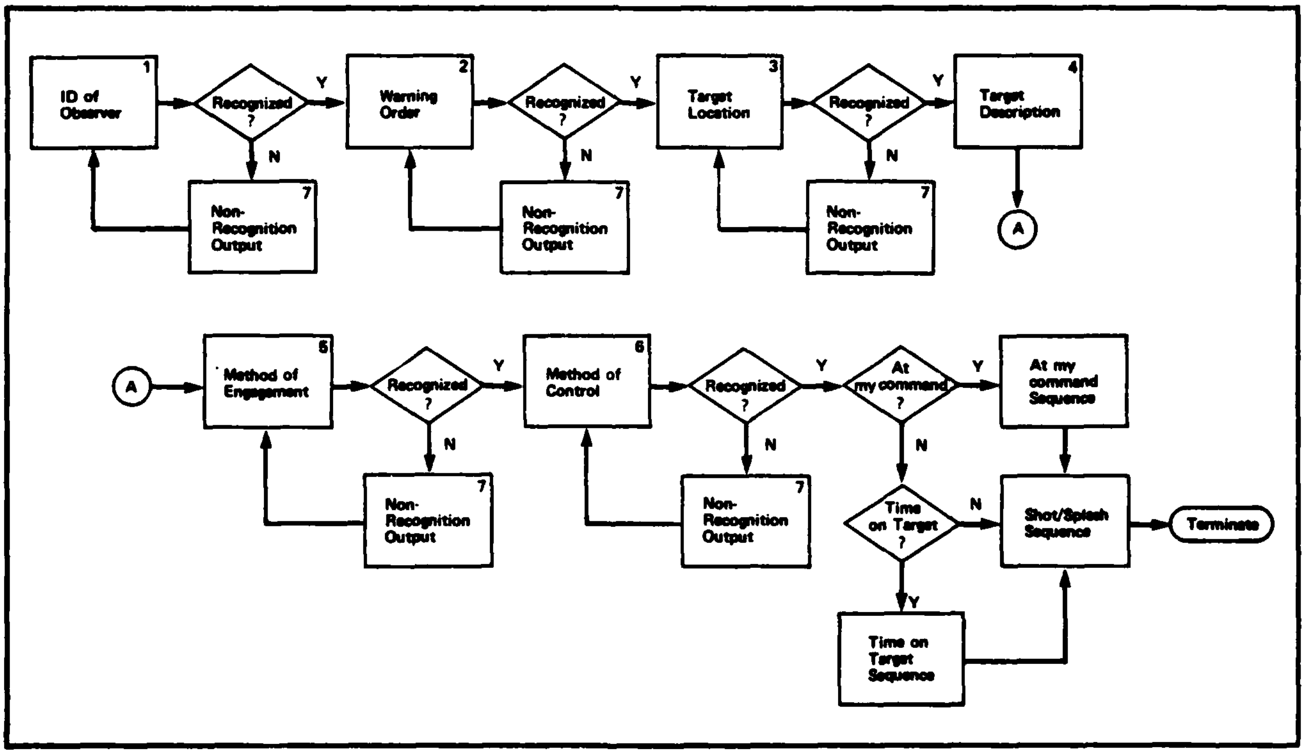


Figure 6-5. **Initial "Adjust Fire" Request (Target of Opportunity)**

Table 6-1

Initial "Adjust Fire" Request
(Reference Figure 6-5)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
"Yankee six five, this is Bravo two six, adjust fire, over." (Symbols 1 and 2)	"Bravo two six, this is Yankee six five, adjust fire, over."
"Grid one, two, three, four, five, six. Direction five, six, zero, zero mils. One, three, zero, zero meters. One armored vehicle in open. Over." (Symbols 3, 4)	"Grid one, two three, four, five, six. Direction, five six, zero, zero mils. One, three, zero, zero meters. One armored vehicle in open. Over."
"One marking round, will adjust. Over." (Symbol 5)	"One marking round, will adjust. Over."
"Roger, Out." TERMINATE OR	
"At my command, Over." or "Time on target, one, four, one five hours. Over." (Symbol 6)	"At your command. Over." or "Time on target one, four, one five hours. Over."
"Roger. Out." TERMINATE	
NOTE: To minimize voice recognition requirements, it is suggested that "At my command" and "time on target" methods of control be omitted. In this manner, the method of control component of the fire mission can be avoided entirely.	
INPUT NOT RECOGNIZED BY SYSTEM. (Symbol 7)	"Bravo two six, this is Yankee six five. You came in broken and distorted. Say again last transmission. Over."

Subsequent "Adjust Fire" Requests

Following the initial "Adjust Fire" request, the system user will provide adjustments to the FDC. The sequencing and logic of this communication are shown in Figure 6-6 in flowchart form. As was the case with the previously described communication sequence, each symbol in this flowchart requiring explanation is noted in Table 6-2.

The subsequent adjust fire request addressed in Figure 6-6 will normally follow the initial adjust fire request. The only time this will not be true is when the initial request is on target (this should occur no more than 50% of the time). When this occurs, the requestor will call for a "Fire for Effect" immediately following his initial adjust fire request which resulted in one marking round impacting. Each subsequent adjust fire request will also result in one marking round impacting. After the marking round is within 50 meters of the target, the requestor will call for a "Fire for Effect" handled in the manner described in Figure 6-7.

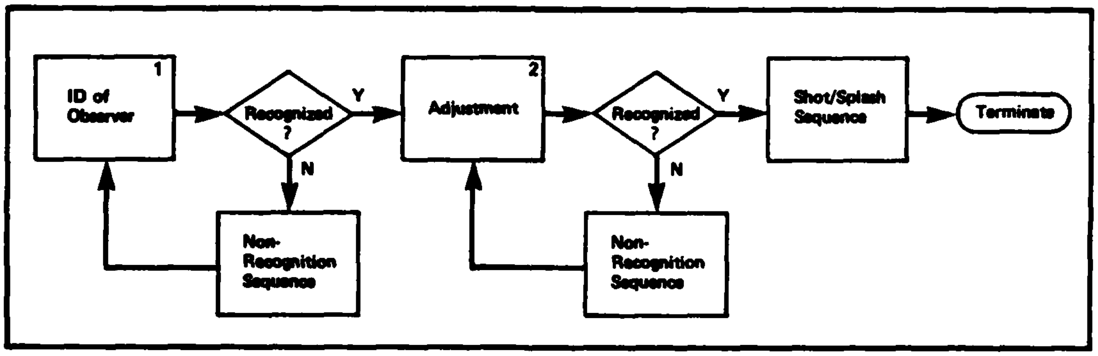


Figure 6-6. **Subsequent "Adjust Fire" Request**

Table 6-2

Subsequent "Adjust Fire" Request
(Reference Figure 6-6)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
"Yankee six five, this is Bravo two six, Add" OR "Drop ___ ___ ___ meters, Left" OR Right ___ ___ ___ meters. Over." (Symbols 1 and 2) "Roger, Out." **NOTE:** Following adjustment and preceeding over, requestor could say "Fire for Effect." If this occurs, the FDC would repeat the mission verbatum. When a fire for effect request is requested, five rounds will impact. If the request is restricted to an adjustment only, only one marking round will impact as was the case with the initial adjust fire request.	"Bravo two six, this is Yankee six five. Add" OR "Drop ___ ___ ___ meters, Left" OR "Right ___ ___ ___ meters. Over."

Fire for Effect Request

A fire for effect request results in a minimum of five (Battery one) to as many as twenty-five (Battery 5) artillery rounds impacting at the requested location. A fire for effect can be requested in a number of ways. Normally, it follows either an initial or subsequent adjust fire request. When this occurs, the system user will merely establish contact with the FDC, and say "Fire for Effect." However, he may give a grid and direction (Initial Adjust Fire) or adjustment from last fire mission (Subsequent Adjust Fire) in the same manner previously defined. The only difference is he will state "Fire for Effect" instead of "Adjust Fire" when initiating the request. It is the former procedure that is outlined in Figure 6-7 and Table 6-3 which, combined, illustrate the communication sequence and basic logic of this communication.

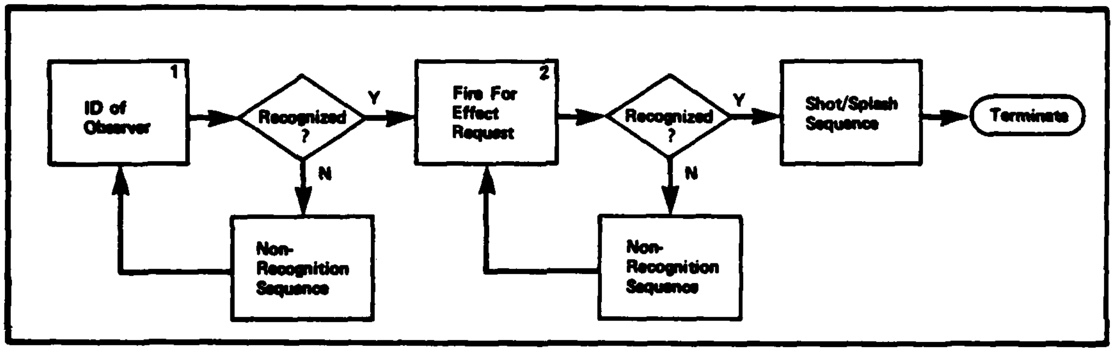


Figure 6-7. **Fire for Effect Request**

Table 6-3

Fire for Effect Request
(Reference Figure 6-7)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
"Yankee six five, this is Bravo two six. Over" (Symbol 1)	
	"Bravo two six, this is Yankee six five. Over." (Symbol 1)
"Yankee six five, this is Bravo two six. Fire for Effect. Over." (Symbol 2)	
	"Bravo two six, this is Yankee six five. Fire for Effect. Over." (Symbol 2)
"Roger, Out."	
OR	
When more than a Battery one is requested:	
"Yankee six five, this is Bravo two six. Fire for Effect, Battery______ Over." (Symbol 2)	
	"Bravo two six this is Yankee six five. Fire for Effect, Battery______ Over." (Symbol 2)
"Roger. Out."	

Shot/Splash Sequence

After the termination of an initial or subsequent adjust fire and fire for effect missions, the FDC will inform the requestor when his mission is fired by the guns (shot) as well as when the rounds should be impacting (normally within five seconds) in the area requested by the observer (splash). This procedure is followed by the FDC for all marking rounds as well as fire for effects (five or more rounds impacting).

The sequencing and basic logic for this communication are shown in Figure 6-8. Unlike the previously discussed communications, this communication is initiated by the system as opposed to the system's user. Each symbol in the flowchart requiring explanation is numbered and explained in Table 6-4 which immediately follows the flowchart. The non-recognition symbol's procedure is identical to that described in and Table 6-1.

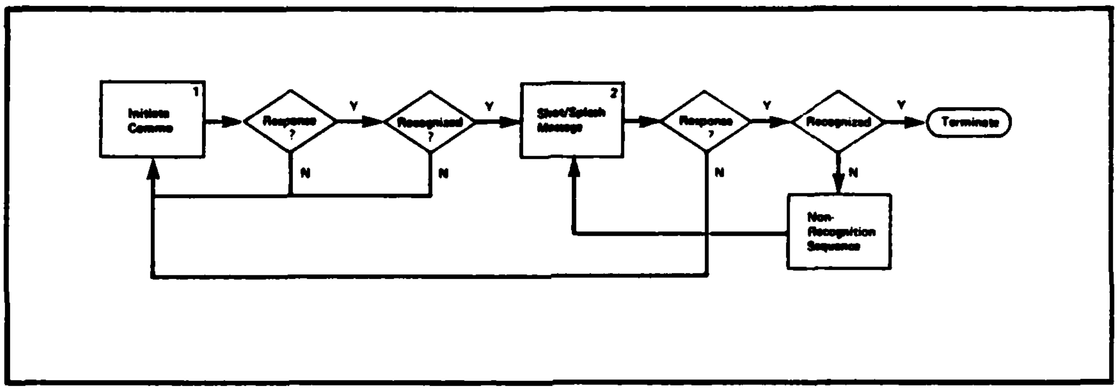


Figure 6-8. **Shot/Splash Sequence (Initiated by System)**

Table 6-4

Shot/Splash Sequence
(Reference Figure 6-8)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
SHOT SEQUENCE	
	"Bravo two six, this is Yankee six five. Shot. Over." (Symbols 1 and 2)
"Yankee six five, this is Bravo two six. Shot. Out" (Symbol 2) OR (Abbreviated response) "Shot. Out." (Symbol 2)	
SPLASH SEQUENCE	
	"Bravo two six this is Yankee six five. Splash. Over." (Symbols 1 and 2)
"Yankee six five, this is Bravo two six. Splash. Out." (Symbol 2) OR (Abbreviated response) "Splash. Out." (Symbol 2)	

Authentication Sequence

At any point during an initial/subsequent adjust fire or fire for effect request, the FDC may choose to exercise an authentication communication security procedure. The sequencing and basic logic for this communication is shown in Figure 6-9.

The broken lines entering the first symbol and exiting the last symbol in this flowchart represent the fact that this communication procedure can occur at any point during an initial or subsequent adjust fire mission. This communication will always be initiated by the FDC (i.e., system). Explanations of the symbols requiring discussion are contained in Table 6-5. All authentication codes will be included in the CEOI provided to the user which he should review before initiating interaction with the system.

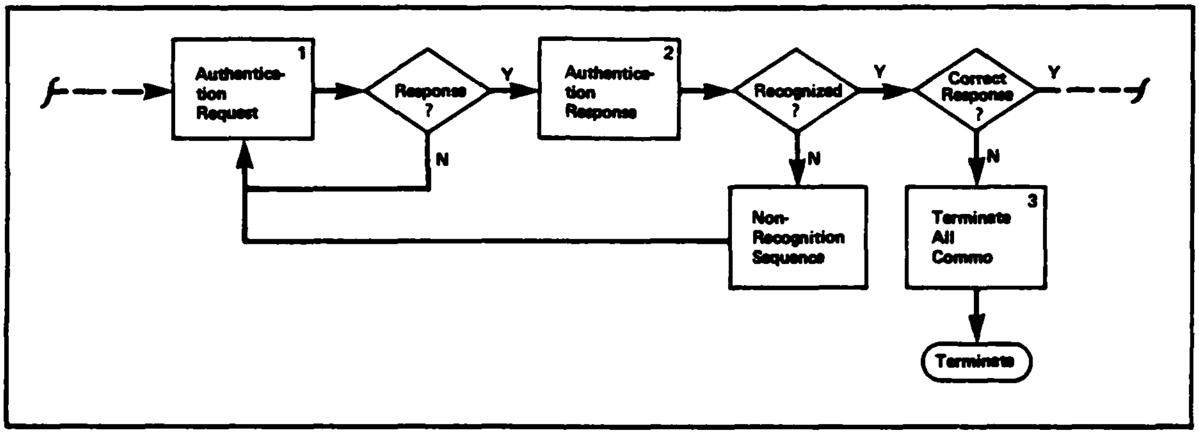


Figure 6-9. **Authentication Sequence**

Table 6-5

Authentication Sequence
(Reference Figure 6-9)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
	". . . Authenticate Alpha, Bravo. Over." (Symbol 1)
"Authentication is Charlie. Over." (Symbol 2)	
	"Roger. . . ." (FDC PROCEEDS WITH INTERRUPTED COMMUNICATION)
IF THE RESPONSE IS INCORRECT OR DOES NOT OCCUR WITHIN TWENTY SECONDS, ALL COMMUNICATIONS WITH THE REQUESTOR WILL TERMINATE IMMEDIATELY. (Symbol 3)	

Jamming Sequence

At any point, during any communication, the enemy can initiate a variety of jamming procedures. These jamming actions manifest themselves over radio nets in a variety of ways depending on the jamming technique employed. When this occurs, friendly force communicators must switch to an alternate frequency if they wish to continue communications. Alternate frequencies are identified in CEOI's which will be provided to the user of system. It will be the users' responsibility to know how to use the CEOI to identify the correct alternate frequency.

Jamming can be initiated by the system at any time. The communication sequencing and basic logic for this are illustrated in Figure 6-10. Any symbols in this flowchart requiring explanation, are explained in Table 6-6 immediately following Figure 6-10.

After the FDC announces he is switching to the alternate frequency (Symbol 2 in Figure 6-10), the snythesizer will continue to output the jamming "noises" until the user "changes frequencies." To change frequencies, the synthesizer will ask the user what the alternate frequency is, the user will verbally respond, and the system will then acknowledge the correct alternate frequency informing the user he can now proceed with his communications.

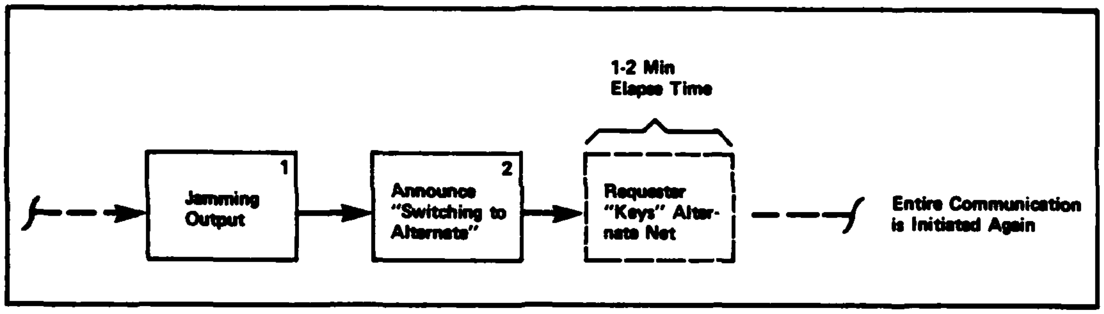


Figure 6-10. Jamming Sequence

Table 6-6

Jamming Sequence
(Reference Figure 6-10)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
	Jamming manifestations described in contract deliverable 0002AG, pages 38 and 39. (Symbol 1)
	After a "lull" in jamming manifestations, the FDC will state ". . . switching to alternate frequency. Out." (Symbol 2)
	"Please identify the correct alternate frequency. You have 90 seconds to respond. If you do not respond within this time limit, the system will disconnect. Your time limit begins now." (In a voice other than that used for the FDC — perhaps a "robotic sounding" female voice).
Hypothetical Correct Response—	
"Three, niner, seven, five."	"That is correct. You must now initiate your last communication with the FDC from the beginning. Thank you." (Same voice as last output)
Hypothetical Incorrect Response—	
"Three, niner, six, eight."	"That is incorrect. You may try again if your time limit has not expired." (Same voice as last output.)
Unrecognizable Response.	"I did not understand what you said. Please repeat the frequency number again. I will add 15 seconds to your time limit." (Same voice as previous output.)
Time Limit Expires.	"I'm sorry, your time is up. That concludes this exercise. You may start all over if you wish. Better luck next time."

Communications Check Sequence

When first entering a radio net or when experiencing problems communicating over the net, one of the members of a communication's net may request a communications or commo check. It is strongly recommended that the system initiate the commo check with the user. An appropriate time to do this would be before any other communications occur over the system, i.e., within seconds after the user gets the "system up." It could be written in the system's procedures that the system will initiate contact with the user and following this initialization procedure, the user is free to communicate with the system as and when he pleases.

Figure 6-11 illustrates the sequence and basic logic of this sequence. Any symbols in the flowchart requiring explanation are numerically numbered and referenced in Table 6-7.

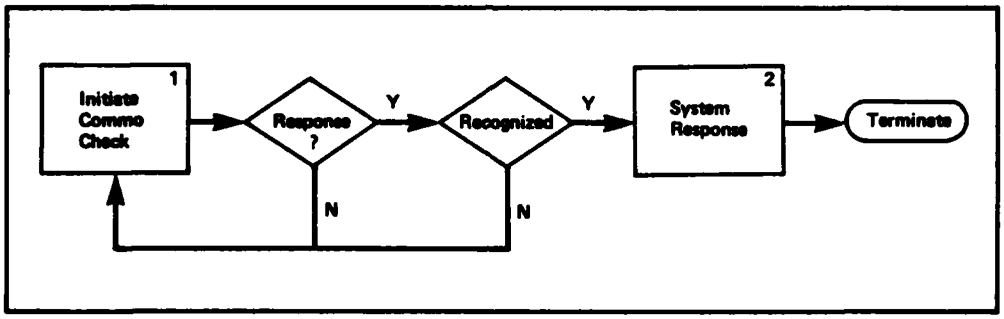


Figure 6-11. **Commo Check Sequence (Initiated by System)**

Table 6-7

Communications Check Sequence (Reference Figure 6-11)

Recognition Requirement (Requester)	Synthesis Requirement (FDC)
	"Bravo two six, this is Yankee six five. Commo Check, Over." (Symbol 1)
"Yankee six five, this is Bravo two six. Lima Charlie" OR "Loud and Clear" then "How me, Over" OR "Over." (Symbol 1)	
	"Hear you same. Out." (Symbol 2)

General and Specific SIMCOMM Procedural Rules

General and specific rules which must be incorporated into the SIMCOMM design if the conceptual design just discussed is to operate effectively will be discussed in this subsection.

These rules can be viewed as an artifact of the SIMCOMM and, in fact they are. However, the rules do not result in the SIMCOMM simulations being unrealistic. Rather, though peculiar to SIMCOMM's operation, they dictate the sequence of realistic communication activities. These rules should not be viewed as exhaustive. As SIMCOMM's development progresses, they will no doubt be modified. However, as of this writing they include:

1. Initial missions will always be "Adjust Fire" missions with the FDC responding with a "marking round."

2. Initial target location will always be reported using a six-digit grid coordinate.

3. Direction will always be given in "MILS."

4. Adjustments will always:
   - Follow initial request unless a "fire for effect" is requested next.
   - Be made in 50 meter increments.
   - "ADD" or "DROP" adjustment given first followed by "Left" or "Right" adjustment.

5. Strict radiotelephone procedures are in effect. Therefore, each communication with the FDC will begin with identification of observer.

6. FDC will always provide "SHOT" and "SPLASH."

7. System will initiate all commo checks.

8. Initial adjust fire mission will result in one marking round.

7 — SYSTEM'S OVERVIEW

The purpose of this section is to detail SIMCOMM's hardware and software configuration as well as the rationale for the configuration engineered. SIMCOMM's hardware/software configuration is sensitive to the general requirements discussed in Section 6, i.e., it is a standalone configuration, portable, can be interfaced with most mini/micro/large main frame computers in the Army's inventory and can be easily replicated given its availability on an off-the-shelf basis and low costs.

HARDWARE

SIMCOMM's hardware will consist of an off-the-shelf Texas Instruments Professional Computer with 256-Kilobytes of dynamic, read/write primary memory, 10-Megabytes secondary memory on a Winchester hard-disk as well as a 320-Kilobyte double sided/double density floppy disk drive. The display will be a high resolution (>18-MHz bandwidth), monochrome video monitor that is capable of displaying a picture with 720 by 320 pixel (picture element) resolution with an 8-level gray scale. Also included will be an RS-232C interface board (for communicating to other computers) and TI's Speech Command board for speech synthesis and recognition.

SOFTWARE

Software supplied will consist of the following packages. The operating system will be MS-DOS from Microsoft and runs on most Intel 8086/8 microcomputers. This situation is accompanied by a large selection of software available from many third party sources.

The major portion of the software development will be done in the language C. This language, developed at Bell Labs, is a structured but very flexible and portable language and was chosen on these merits. The compiler to be used is the Lattice C Compiler and was chosen on its ability to generate very fast, efficient machine instructions for the C programming language.

For accessing the TIPC's hardware functions more efficiently and for portions of SIMCOMM that require maximum speed, assembly language is needed. Assembly language is for programming in the microprocessor's "native language". The assembler to be used will be Microsoft's Macro Assembler chosen for its macro capabilities and the software tools that come with it. These tools will greatly help in combining all of the C and assembly language programs together into a working software system.

A text editor will be necessary to create both the programs and general text associated with SIMCOMM. Although a text editor comes with the MS-DOS package it is line oriented and cumbersome to use. The editor chosen for the TIPC is PMATE, from Aox Incorporated, which is a full screen text editor, as opposed to a line editor in that a full screen of text (22-lines) is displayed at one time instead of a single line.

Two other programming languages from Microsoft may be included with SIMCOMM but, at this point in time, are not needed in its development. The languages are MS-BASIC and MS-FORTRAN. Both are standards for usage under MS-DOS as well as very popular software products.

OFF-THE-SHELF/FLEXIBILITY/ MODULARITY/EXPANDABILITY

The SIMCOMM system will be produced entirely from off-the-shelf products (hardware and software) available through Texas Instruments. A large selection of software is on the market for MS-DOS and the TIPC through third party software vendors as well as TI for further SIMCOMM development by ARI.

Due to C's modular nature as a programming language, SIMCOMM's software can be implemented rather easily from a top-down, structured approach. C's modularity comes from its programs being a collection of programming modules or functions. When done correctly this collection of functions can be extremely flexible and open to expandability after the design and implementation stages. Most changes can usually be handled by the proper insertion/deletion of appropriate functions.

8 — SYSTEM'S DEVELOPMENT, TESTING AND EXPANSION

The objective of this research is to advance the application of voice technology to simulation and training devices. Two research phases are required to achieve this objective. The purpose of the first phase was to develop the conceptual design of an interactive voice SMI and its associated linguistic database. The objectives of Phase I have been achieved with the publication of this technical report.

Phase II requires completion of three tasks: (1) system's design; (2) build and test the system, and; (3) develop a future expansion and integration plan. The purpose of this section is to present the approach to each of these tasks.

SYSTEM'S DESIGN

The objective of this task is to develop design documentation from which SIMCOMM will be built. A sound departure point and foundation for this task has been established with: a survey and evaluation of state-of-the-art voice synthesis and recognition technologies (Section 3); a comprehensive definition of tactical voice communications (Section 4); a theoretical construct of tactical voice communications (Section 5) which can serve as the framework for any tactical voice communications linguistic database, and; the conceptual design of a voice interactive system (Section 6) which addresses the system's functional requirements, internal/external interfaces and progression/sequencing through subsystems.

To accomplish the objective of this task, three steps will be required and performed in the following sequence:

- Modify Conceptual Design — The feasibility of satisfying SIMCOMM's conceptual design requirements has been determined, and a hardware/software configuration conceived (Section 6). However, SIMCOMM's conceptual design is being documented for the first time in this report, so review by other organizations may be necessary. These organizations might include other ARI laboratories not involved in Phase I and TRADOC organizations, such as the Artillery School. Resulting comments/recommendations will be considered, and SIMCOMM's conceptual design modified accordingly during this step.
- Develop Preliminary Design — During this step, the call for fire, COMSEC, and jamming modules (specified in Section 6) will be verified in terms of correct terminology, procedures, and sequencing. This will be accomplished with the assistance of appropriate military authorities (e.g., the Artillery School). A preliminary design of the SIMCOMM will then be prepared in a format similar to Section 6. This will be reviewed by ARI, TI, and HumRRO, and modified as required.
- Develop Detailed Design — In this step, specific software logic will be designed and documented, visual presentation requirements identified, sequencing/interaction requirements (i.e., between voice synthesis/recognition and visual stimulus) identified and supporting material requirements (e.g., user documentation, CEOIs, topographical maps) specified. Each of these will be documented and presented to ARI for approval.

BUILD AND TEST SIMCOMM

The objective of this task is to build and test the SIMCOMM. Accomplishing this task will require three steps performed in the sequence presented below:

- System Development and Test Planning — During this step the SIMCOMM will be developed by software/engineer specialists. The necessary hardware will be procured and configured, and all necessary software will be coded. Plans for SIMCOMM's testing will be developed and submitted to ARI for approval. The test plans will specify all troop support requirements, data collection, and analysis procedures.
- Pilot Test — There will be two SIMCOMM pilot tests. The first of these will involve demonstrating the system to Artillery School personnel. They will evaluate SIMCOMM's technical accuracy. The second pilot test will involve demonstrating the SIMCOMM to a few personnel from a combat arms unit. The latter pilot test will center on user acceptance. SIMCOMM will be modified as required based upon the results of these pilot tests.
- Field Test — A major field test will be conducted using personnel from the combat arms. The evaluation or testing criteria for this field test will center around the degree to which voice technology enhances soldier performance in an operational and/or training systems. In an operational environment, SIMCOMM could be used as a job aid providing the soldier with intelligence about the disposition of enemy forces (e.g., an abbreviated VINT[2]). In this case, the soldier's proficiency in employing indirect fire with (i.e., experimental group) and without (i.e., control group) this intelligence could be evaluated. To evaluate the effects of voice technology on training effectiveness, soldiers could be taught call for fire procedures using the SIMCOMM (experimental group) or in a more traditional manner not using the SIMCOMM (control group). Factors such as time to teach, costs, and performance could then be compared. The results of the field test(s) will be formally documented and presented to ARI for review.

EXPANSION AND INTEGRATION PLAN

The purpose of this task is twofold. First, a plan for expanding SIMCOMM's capabilities will be developed. Second, the feasibility of integrating SIMCOMM (in its original or a modified form) into existing Army training and/or operational systems will be investigated. SIMCOMM's integration with existing systems will be facilitated by its modular design and incorporation of a standard RS232 computer-to-computer interface.

9 — SUMMARY

The dramatic introduction of technology onto the battlefield has intensified in recent years and is expected to continue at an accelerated rate. This results in increased human performance demands. Human performance is physiologically limited in terms of the rate at which a human can process information and physically perform tasks. Though there is little that can be done about the human's physiological limitations, human performance capabilities can be improved in other ways. Human performance capabilities of Army personnel can best be improved by either raising enlistment standards (e.g., education, aptitude) and/or providing more and better training. Enlistment standards are expected to remain relatively constant during the foreseeable future. Though Army training has improved and will continue to do so, the training environment is burdened with everincreasing demands (to a large extent attributable to the proliferation of technology on the battlefield) while simultaneously experiencing diminishing resources (e.g., time and money). As a result, human performance capabilities in the Army are also expected to remain relatively constant.

Given a relatively constant soldier performance capability and increased complexity of the war machines which the soldier must operate, a serious problem surfaces. Today's soldier is required to do more than merely point his rifle. He is expected to operate complex, technologically sophisticated systems upon which victory on the battlefield is highly dependent. One way to lessen the gap between soldier performance capabilities and technologically advanced systems is to improve the soldier/machine interface (SMI).

The SMI can be improved in a variety of ways including environmental/job design, artificial intelligence, decision aids and voice recognition/synthesis. The objective of the research address in this technical report is to advance the application of evolving speech technology to Army tactical operational and training systems. To accomplish this, a review of the state-of-the-art of voice technology was performed and its potential benefits to tactical operational and training systems' SMIs determined. A definition and taxonomy of tactical voice communications were developed. A conceptual design of a computer-assisted simulation of tactical voice communications (SIMCOMM) was then developed which includes its voice interactive protocols, i.e., speech synthesis and recognition requirements, and a specification of its hardware configuration.

The SIMCOMM is a standalone, portable system which will demonstrate voice synthesis/recognition and, to a lesser degree, artificial intelligence technologies. Its associated hardware/software configuration is compatible with most mini/micro and large mainframe computers currently used in the Army. SIMCOMM's software is modular by design. Given these attributes, SIMCOMM can be transported for demonstration purposes, interfaced to existing tactical operational and training systems, and its speech synthesis/recognition capabilities easily expanded. In addition, SIMCOMM can be used as a research tool for investigating voice input/output soldier/machine interfaces. Finally, SIMCOMM's cost is minimal and, as such, facilitates its replication. Combined, SIMCOMM's attributes and capabilities will enable it to advance the application of voice technology in Army environments.

Accomplishing the activities covered in this report constitutes completion of Phase I of this research effort. Phase II, which will be initiated following publication of this report, will encompass the actual building and testing of the SIMCOMM.

REFERENCES

Allen, J. (1976), Synthesis of speech from unrestricted text. *Proceedings of the IEEE,* 64(4), 433-442.

Bales, F.R. (1950), *Interaction process analysis.* Cambridge, MA: Addison Wesley Press.

Born, T.B. (September 1981), *Empirical analysis of systematic communication errors.* Monterey, CA: Naval Postgraduate School, (DTIC No. AD A109573).

Briggs, G.E. & Johnston, W.A. (June 1967), *Team training.* (Technical Report 1327-4) Orlando, FL: U.S. Naval Training Device Center.

Brown, R.L. (June 1967), *A content analysis of communications within Army small-unit patrolling operations.* (HumRRO Technical Report 67-7) Alexandria, VA: Human Resources Research Organization.

Clarke, E.T., Baicker, J.A., Cox, J.W., Kay B.R.R., Clement, J.B., & Benson, J.W. (1953), *A study of combat communications: Korea, January - July, 1952.* (Technical Memorandum ORO-T-43 FEC). Chevy Chase, MD: Operations Research Office, Johns Hopkins University.

Department of the Army. (25 July 1980), *Armored and mechanized brigade operations.* Field Manual 71-3, Washington.

Department of the Army. (30 September 1977), *The mechanized infantry platoon/squad.* Field Manual 7-7, Washington.

Department of the Army. (30 June 1977), *The tank and mechanized infantry company team.* Field Manual 71-1, Washington.

Department of the Army. (31 March 1977), *Combat communications within the division.* Field Manual 11-50, Washington.

Department of the Army (31 December 1976), *The rifle squads (mechanized and light infantry).* TC 7-1, Washington.

Doddington, G.D., & Schalk, T.B. (1981), Speech recognition: turning theory to practice. *IEEE Spectrum,* 18(9), 26-32.

Federman, P. & Siegel, A.I. (1965), *Communications as a measurable index of team behavior.* (NAVTRADEVCEN 1537-1) Port Washington, NY: U.S. Naval Training Device Center.

Gagnon, R.T. (April 1978), Votrax real-time hardware for phoneme synthesis of speech. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing,* Tulsa, Oklahoma.

Galbraith, J.R. (May 1974), Organizational design: An information processing view. *Interfaces,* 4(3).

Hankin, B.D. (September 1952), *A new approach to the communication problem presented by division signals,* (Report No. 31). England: Army Operational Research Group.

Hankin, B.D. & Love, L.G. (July 1952), *An analysis of the Signal Services at Army Group "A" as carriers of "Information" on Exercise GRANDE ALLIANCE.* (Report No. 28). England: Army Operational Research Group.

Hannaman, D.L., Underhill, W.G., Laurence, M.T. & Chambers, R.M. (September 1982), *Behavior simulation: Definition, field observation and automation of small unit leader tactical behavior.* (HumRRO Final Report 82-17). Alexandria, VA: Human Resources Research Organization.

Hazell, J.W. & Leyzorek, M. (1953), *A preliminary study of combat information-handling* (Technical Memorandum ORO-T-226). Chevy Chase, MD: Operations Research Office, Johns Hopkins University.

Itakura, F. (1975), *Minimum prediction residual principle applied to speech recognition.* "IEEE Transactions on Acoustics, Speech, and Signal Processing," Volume ASSP-23, 67-72.

Kelly, M.J. & Chapanis, A. (1977), Limited vocabulary natural language dialogue. *International Journal of Man-Machine Studies,* 9, 479-501.

Krumm, R.L. & Farina, A.J. (February 1962), *Effectiveness of integrated flight simulator training in promoting B-52 crew coordination.* (MRL Tech Documentary Report 62-1). Wright Patterson AFB, OH: Aerospace Medical Research Laboratories.

Kurzweil, R. (Spring 1979), Kurzweil talking terminal announced; first unlimited vocabulary. *The Kurzweil Report.* Kurzweil Computer Products, Cambridge, Massachusetts.

Ladefoged, P. (1975), *A course in phonetics.* New York: Harcourt Brace Jovanovich.

Lennahan, C.B. (December 1960), *Communications data for tank battalions: Ft. Stewart, Ga.* (Technical Paper ORO-TP-13). Bethesda, MD: Operations Research Office, Johns Hopkins University.

Lin, N., Goudie, G., Frantz, G., & Brantingham, L. (November 1980), Text to speech using LPC allophone stringing. *Proceedings of the Fall IEEE Consumer Electronics Conference,* Chicago.

Lin, N. (1973), *The study of human communication.* Bobbs Merrill Co.

Markel, J.D., & Gray, A.H. (1976), *Linear prediction of speech.* New York: Springer-Verlag.

Michaelis, P.R. (October 1981), Choosing a vocabulary for a speech synthesizer. *Proceedings of the National Electronics Conference,* Chicago, 239-243.

Michaelis, P.R., Chapanis, A., Weeks, G.D., & Kelly, M.J. (1977), Word usage in interactive dialog with restricted and unrestricted vocabularies. *IEEE Transactions on Professional Communication*, PC-20, 214-221.

Michaelis, P.R., & Wiggins, R.H. (1982), A human factors engineer's introduction to speech synthesizers. In A. Badre & B. Shneiderman (Eds.), *Directions in human-computer interaction.* Norwood, New Jersey: Ablex Publishing Corporation.

Root, R.T., Epstein, K.I., Steinheiser, F.H., Hayes, J.F., Wood, S.E., Sulzen, R.H., Burgess, G.G., Mirabella, A., Erwin, D.E., & Johnson, E. (October 1976), *Initial validation of REALTRAIN with Army Combat Units in Europe.* (ARI Report No. 1191). Arlington, VA: Army Research Institute.

Ruark, A.E. (May 1951), *Future of naval communications.* (PRU-145) Baltimore, MD: Programs Research Unit, Institute for Cooperative Research, Johns Hopkins University.

Schalk, T.B., & McMahon, M. (May 1982) "Low cost word recognition using programmable VLSI." Paper presented at N.A.E.C.O.N. Conference, Dayton, Ohio.

Stein, E.S. & Bruce, J.M. (June 1981), *The relationship of communications to combat unit performance.* (Draft Technical Report) Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Tremain, T.E. (1982) The government standard linear predictive coding algorithm: LPC-10. *Speech Technology* 1(2), 40-49.

Wiggins, R.H., & Brantingham, L. (August 1978), Three-chip system synthesizes human speech. *Electronics*, 51(18), 109-116.

Woods, D.L. (1974), *A history of tactical communications.* New York, Arno Press.

BIBLIOGRAPHY

Ainsworth, L.L., & Bishop, H.P. (July 1971) *The effects of a 48-hour period of sustained field activity on tank crew performance* (HumRRO Technical Report 71-16). Alexandria, VA: Human Resources Research Organization.

Barr, A., Beard, M., & Atkinson, R. (1976) The computer as a tutorial laboratory: The Stanford BIP project. *International Journal of Man-Machine Studies*, 8, 567-596.

Baum, D.R., Modrick, J.A., Hollingsworth, S.R. (April 1982) *Team training for command and control systems: Recommendations for application of current technology* (AFHRL Technical Paper 82-9). Wright-Patterson Air Force Base, OH: Air Force Human Resources Laboratory, (DTIC No. AD-A-113-768).

Brick, D.B., & Ellersick, F.W. (February 1981) *Challenges and opportunities in Air Force tactical communications* (ESD Technical Report 81-115). Hanscom Air Force Base, MA: Electronic Systems Division, (DTIC No. AD-A097677).

Brown, F.L. (April 1967) *Critical combat performances, knowledges, and skills required of the infantry rifle platoon leader: Use of indirect supporting fires* (Research By-Product). Alexandria, VA: Human Resources Research Organization.

Brown, F.L. (June 1967) *A content analysis of communications within Army small-unit patrolling operations* (HumRRO Technical Report 67-7). Alexandria, VA: Human Resources Research Organization.

Brown, F.L., & Jacobs, T.O. (April 1970) *Developing the critical combat performance required of the infantry rifle platoon leader* (HumRRO Technical Report 70-6). Alexandria, VA: Human Resources Research Organization.

Brown, F.L., & Magner, G.J. (June 1969) *Critical combat performances, knowledges, and skills required of the infantry rifle squad leader: Defensive operations* (Research By-Product). Alexandria, VA: Human Resources Research Organization.

Carroll, J.B. (1967) On sampling from a lognormal model of word frequency distribution. In H. Kucera and W.N. Francis (Eds.), *Computational analysis of present-day American English.* Providence, Rhode Island: Brown University Press.

CG, USARV. (1968) *USARV seminar report: Night operations in RVN (U).* Washington, DC: OACSFOR, Department of the Army.

Chapanis, A., Ochsman, R.B., Parris, R.N., & Weeks, G.D. (1972) Studies in interactive communication: I. The effects of four communication modes on the behavior of teams during cooperative problem-solving. *Human Factors*, 14, 487-509.

Chapanis, A. (1965), *Man-machine engineering.* Belmont, CA: Brooks/Cole.

Chen, Y. (May 1982), An analysis-by-synthesis approach for automatic time segmentation of speech signals. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing,* Paris, France.

Cleary, F.K., & Kelly, H.E. (July 1968) *Critical combat performances, knowledges, and skills required of the infantry rifle platoon leader* (Research-By-Product). Alexandria, VA: Human Resources Research Organization.

Condon, E.V. (1928) Statistics of vocabulary. *Science,* 67, 300.

Controller General of the United States. (January 1980) *The joint tactical information distribution system—How important is it?* (Report to the Congress). Washington, DC: General Accounting Office.

Department of the Army. (No date) *Army training* Army Regulation 350-1. Washington.

Department of the Army. (No date) *Communications-electronics operation instruction (CEOI) program, policy, responsibility and guidance,* Army Regulation 105-64. Washington.

Department of the Army (No date) *Communications security (COMSEC) operations,* Army Regulation 530-2. Washington.

Department of the Army. (No date) *Communications training—sections and platoons* (ASUBJSCD 6-5). Washington.

Department of the Army. (No date) *Integrated tactical communications systems (INTACS) steering committee,* Army Regulation 15-23. Washington.

Department of the Army. (No date) *Opposing force (OPFOR) program,* Army Regulation 350-2. Washington.

Department of the Army (March 1981) *JINTAGGS technical interface design plan (test edition). Appendix B—Bibliography and References.* Washington, DC: Office of the Assistant Chief of Staff for Automation and Communications.

Department of the Army. (1982) *How to conduct training exercises.* Field Manual 25-4. Washington.

Department of the Army. (1982) *How to conduct training in units.* Field Manual 25-3. Washington.

Department of the Army. (June 1982) *How to manage training in units.* Field Manual 25-2. Washington.

Department of the Army (1982) *Training.* Field Manual 25-1. Washington.

Doddington, G.D., & Schalk, T.B. (1981) Speech recognition: Turning theory to practice. *IEEE Spectrum,* 18(9), 26-32.

Egbert, R.L., Katter, R.V., & Greer, G.D. (October 1953) *Incidental observations gathered during research in combat units* (Information Report). Alexandria, VA: Human Resources Research Organization.

Field, K.S. (September 1966) *Marine tactical command and control study (MTACCS)—Annex E: Operations* (Final Report, Vol. VI). Menlo Park, CA: Stanford Research Institute.

Fisher, W., & Goudie, K. (October 1981) *Phonetic correlates of stress and syllabification.* Paper presented at the Annual Meeting of the Michigan Linguistic Society.

Fitzgerald, COL J.P. (April 1982) *Integrated information processing/communications: The key to command, control and intelligence on the extended battlefield.* Carlisle Barracks, PA: U.S. Army War College, (DTIC No. AD A116697).

Ford, W.R., Weeks, G.D., & Chapanis, A. (1980), The effect of self-imposed brevity on the structure of dyadic communication. *The Journal of Psychology,* 104, 87-103.

Frantz, G., & Wiggins, R.H. (March 1981), The development of "solid state speech" technology at Texas Instruments. *The Newsletter of the IEEE Acoustics, Speech, and Signal Processing Society,* 53.

Gabelman, I.J. (Ed.), (April 1979) Techniques for data handling in tactical systems II. *Proceedings of the North Atlantic Treaty Organization Advisory Group for Aerospace Research and Development* (AGARD-CP-251). Neuilly Sur Seine, France: AGARD.

Gabriel, R. & Moriarty, J. (January 1983) Israel's main battle tank pounds its message home. *Military Electronics/Countermeasures,* pp. 124-132.

Gass, W.K. *The TMS-32010 provides speech I/O for the personal computer.* Paper presented at Electro/83.

Germans, J.E., & Baker, J.D. (July 1980) *Embedded training: Utilization of tactical computers to train tactical computer operators* (Technical Report 452). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gossett, T.D., & Velligan, F.A. (December 1982) The Aquila: A versatile, cost-effective military tool shows its potential. *Military Electronics/Countermeasures.* pp. 74-78.

Gottlieb, I.M. (December 1981) From smoke signals to spread spectrum, military communications decides who wins the war. *Military Electronics/Countermeasures,* pp. 42-47.

Government Accounting Office. (October 1981) *The world wide military command and control information system—problems in information resource management.* Washington DC: Author.

Grace, J.A. (March 1980) *A study of some human factors and their impact on command and control systems.* Monterey, CA: Naval Postgraduate School (DTIC No. AD A092703).

Hammett, J.C. (June 1979) *Tactical communications—a precious resource* (Student Essay). Carlisle Barracks, PA: U.S. Army War College.

Hendler, J.A., & Phillips, B. (August 1981) Controlling parsing by passing messages. *Proceedings of the Annual Conference of the Cognitive Science Society.* Berkeley, CA: pp. 228-230.

Henriksen, K., Hannaman, D.L., Olmstead, J.A., Stein, E.S., & Bruce, J.M. (February 1981) *Development of low cost combat unit leader training techniques: Tactical problem solving and decision making* (HumRRO Final Report 81-3). Alexandria, VA: Human Resources Research Organization.

Henriksen, K.F., Jones, D.R., Hannaman, D.L., Wylie, P.B., & Shriver, E.L. (January 1980) *Identification of combat unit leader skills and leader-group interaction processes* (Technical Report 440). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

HumRRO (undated) *List of activities of infantry rifle platoon leaders.* Alexandria, VA: Human Resources Research Organization.

Ingram, John C. (December 1980) *A detailed review of the TACWAR model.* Adelphi, MD: Harry Diamond Laboratories.

Johnson, J.E. (ed.) (1944), *Basic English.* New York: The W.H. Wilson Company.

Knerr, C.S., Barton, H.D., & Lombardo, J.F. (February 1978) *Evaluation instruments for the basic non-commissioned officer course for combat arms soldiers* (Research Problem Review 77-9). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kucera, H., & Francis, W.N. (1967), *Computational analysis of present-day American English.* Providence, Rhode Island: Brown University Press.

Lahey, G.F., & Slough, D.A. (January 1982) *Relationships between communication variables and scores in team training exercises* (NPRDC Technical Report 82-25). San Diego, CA: Navy Personnel Research and Development Center.

Lindsay, R.K. (1963), Inferential memory as the basis of machines which understand natural language. In E.A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill.

Loo, Bradford G. (September 1981) *Simulation of dismounted infantry combat in urban terrain.* Monterey, CA: Naval Postgraduate School (DTIC No. AD A109550).

Madni, A., Steeb, R., & Rubio, D. (June 1978) *Automated computer support of C3 functions in small tactical units: Part 1—Adaptive information selection* (PTR Semi-Annual Technical Report 1061-78-6(1)). Woodland Hills, CA: Perceptronics.

Magner, G. (July 1968) *Critical combat performances, knowledges, and skills required of the infantry rifle platoon leader: Defensive operations* (Research By-Product). Alexandria, VA: Human Resources Research Organization.

Markel, J.D., & Gray, A.H. (1976) *Linear prediction of speech.* New York: Springer-Verlag.

Martin, H.V. (January 1983) Airland battle 2000 is being implemented with a high technology light division. *Military Electronics/Countermeasures,* pp. 28-36.

Maughmer, R.W. (October 1971) *The design and development of a position and azimuth determining system (PADS).* Fort Belvoir, VA: U.S. Army Engineer Topographic Laboratories.

McCommons, R.B., & Inman, L.J. (December 1969) *A preliminary study of variables affecting call processing efficiency in a tactical environment.* Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratories.

McCormick, E.J. (1976), *Human factors in engineering and design* (4th ed.). New York: McGraw-Hill.

McKay, J.B., Cianci, S., Hall, C.E., & Taylor, J.E. (April 1959) *Some factors which have contributed to both successful and unsuccessful American infantry small-unit actions* (Research Memorandum No. 13). Alexandria, VA: Human Resources Research Organization.

Meliza, L.L., Scott, T.D., & Epstein, K.I. (March 1979) *REALTRAIN validation for rifle squads II: Tactical performance* (ARI Research Report 1203). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Michaelis, P.R. (1980) Cooperative problem solving by like- and mixed-sex teams in a teletypewriter mode with unlimited, self-limited, introduced and anonymous conditions. *JSAS Catalog of Selected Documents in Psychology* 10, 35-36 (MS. No. 2066).

Michaelis, P.R. (October 1981) Choosing a vocabulary for a speech synthesizer. *Proceedings of the National Electronics Conference,* Chicago, IL: pp. 239-243.

Michaelis, P.R., & Chapanis, A. (August 1977) *How many words do you need to solve problems?* Paper presented at the 85th Annual Convention of the American Psychological Association, San Francisco.

Michaelis, P.R., Chapanis, A., Weeks, G.D., & Kelly, M.J. (1977) Word usage in interactive dialogue with restricted and unrestricted vocabularies. *IEEE Transactions on Professional Communication,* 20, 214-221.

Michaelis, P.R., Miller, M.L., & Hendler, J.A. (1982) Artificial intelligence and human factors engineering: A necessary synergism in the interface of the future. In A. Badre & B. Shneiderman (Eds.), *Directions in human-computer interaction.* Norwood, New Jersey: Ablex Publishing Corporation.

Michaelis, P.R., & Wiggins, J.H. (1982) A human factors engineer's introduction to speech synthesizers. In A. Badre & B. Shneiderman (Eds.), *Directions in human-computer interaction.* Norwood, New Jersey: Ablex Publishing Corporation.

Miller, E.E., & Bachta, J.M. (July 1978) *An experimental evaluation of a tactical game for company level training* (ARI Technical Report TR78-B12). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Miller, J.R., Kehler, T.P., Michaelis, P.R., & Murray, W.R. (March 1982) *Intelligent tutoring for programming tasks: Using plan analysis to generate better hints.* (Report ONR-TR-82-0818F). Dallas, TX: Texas Instrument Central Research Laboratories.

Ogden, C.K. (1933), *Basic English: an introduction with rules and grammar* (4th ed.). London: Kegan Paul, Trench, Trubner & Co.

Olmstead, J.A., Hannaman, D.L. Kraemer, A.J., Elder, B.L., Henriksen, K., & McConnell, J.N. (1981) *Development of a simulation-based tactical performance and training developments research facility, Vols. 1-4* (HumRRO Report FR-ETSD-81-5). Alexandria, VA: Human Resources Research Organization.

Paolicelli, J., & Griese, W.F. (December 1980) *Interactive communications systems simulation model (ICSSM)* (RADC Technical Report 80-373). Griffiss Air Force Base, NY: Rome Air Development Center (DCLF).

Pierrehumbert, J. (October 1981), Synthesizing intonation. *Journal of the Acoustical Society of America.* 70(4), 985-995.)

Poock, G.K. (April 1980) *Using voice input to operate a distributed computer network.* Monterey, CA: Naval Postgraduate School.

Rabiner, L.R., & Shafer, R.W. (1978), Digital processing of speech signals. Englewood Cliffs, New Jersey: Prentice-Hall.

Root, R.T., Hayes, J.F., Word, L.E., Shriver, E.L., & Griffin, G.R. (July 1979) *Field test of techniques for tactical training of junior leaders in infantry units: Project EFFTRAIN)* (ARI Technical Report 79-A21). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Saleh, J., & Crolotte, (June 1978) *Automated computer support of C3 functions in small tactical units: Part II—Adaptive decision aiding* (PTR Semi-Annual Report 1061078-6 (II)). Woodland Hills, CA: Perceptronics.

Samet, M.G., & Geiselman, R.E. (May 1980) *Guideline development for summarization of tactical data* (Technical Report 458). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Scott, T.D., Meliza, L.L., Hardy, G.D., Banks, J.H., & Word, L.E. (March 1979) *REALTRAIN validation for armor/anti-armor teams* (ARI Research Report 1204). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Shank, E.G., McJilton, W.N., & Mullen, W.C. (July 1981) *Communications tests with tactical radios in large urban area.* Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratories.

Shavers, V.T. (April 1982) *Command, control, communications (CS) as a tactical force multiplier-myth or reality* (Student Essay). Carlisle Barracks, PA: U.S. Army War College.

Shirley, J.P., Flowers, W.F., Swanson, R.S., Ryan, R.B., & Crair, M.L. (September 1966) *The management and use of tactical radio frequencies in the Republic of Vietnam (U)* (Special Report No. 1). Booz Allen Applied Research Inc., Southeast Asia Operations, (DTIC No. AD-378-779).

Shriver, E.L., Griffin, G.R., Hannaman, D.L., & Jones, D.R. (October 1974) *EFFTRAIN research evaluation conducted at Fort Lewis, Washington.* Unpublished manuscript, (Available from the U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333).

Shriver, E.L., Griffin, G.R., Hannaman, D.L., & Jones, D.R. (August 1975) *Development of training test situations for small units—field tryout of the EFFTRAIN training model in the Berlin Brigade.* Unpublished manuscript. (Available from the U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333).

Smith, M. (January 1983) New tactics, new weapons will derive from the Lebanon conflict. *Military Electronics/Countermeasures,* pp. 116-121.

Songco, D.C., Allen, S.I., Plexico, P.S., & Morford, R.A. (May 1980), How computers talk to the blind. *IEEE Spectrum,* 34-38.

Steck, T.T. (October 1980) *Camouflage and deception techniques for urban warfare* (MERADCOM Report 2310). Fort Belvoir, VA: U.S. Army Mobility Equipment Research and Development Command.

Tactical Command and Control Division, & Keirsey, D.M. (February 1980) *Natural language processing applied to Navy tactical messages* (NOSC Technical Document 324). San Diego, CA: Naval Ocean Systems Center.

Taylor, J.E., McKay, J.B., Hall, C.E., & Cianci, S.N. (April 1959) *PLATTRAIN premises and training implications related to improving the tactical proficiency of rifle platoons* (Research Memorandum No. 12). Alexandria, VA: Human Resources Research Organization.

Thomas, G. (December 1982) Soviets' fight-to-win doctrine incorporates radio electronic combat. *Military Electronics/Countermeasures,* pp. 36-41.

Thrall, R.M. (Ed.). (1976) *Proceedings of the workshop on decision information for tactical command and control* (AFOSR Technical Report 78-0648). Bolling Air Force Base, Washington, DC: Air Force Office of Scientific Research.

USAARMS. (June 1977) Tank and mechanized infantry company team (how to fight). Field Manual 71-1 (HTF). Washington, DC: Department of the Army.

USAARMS. (June 1977) The tank and mechanized infantry battalion task force (how to fight). Field Manual 71-2 (HTF). Washington, DC: Department of the Army.

USAARMS. (July 1977) Cavalry (how to fight). Field Manual 17-95 (HTF). Washington, DC: Department of the Army.

USACAC. (August 1982) Operations (how to fight). Field Manual 100-5 (HTF). Washington, DC: Department of the Army.

USACC. (October 1966) Radio frequency management. Field Manual 24-2. Washington, DC: Department of the Army.

U.S. Army Nuclear Agency (August 1976) *Nuclear blackout of tactical communication* (Nuclear Notes No. 4). Fort Bliss, TX: U.S. Army Nuclear Agency, (DTIC No. AD-A044561).

USAIS. (November 1975) How to prepare and conduct military training. Field Manual 21-6. Washington, DC: Department of the Army.

USAIS (September 1977) The mechanized infantry platoon and squad (how to fight). Field Manual 7-7 (HTF). Washington, DC: Department of the Army.

USAIS. (April 1978) The infantry battalion (how to fight). Field Manual 7-20 (HTF). Washington, DC: Department of the Army.

USAIS. (December 1980) The infantry platoon and squad (how to fight). Field Manual 7-8 (HTF). Washington, DC: Department of the Army.

USAIS. (January 1982) The infantry platoon and squad (how to fight). Field Manual 7-10 (HTF). Washington, DC: Department of the Army.

USASC&FG. (July 1965) Field radio techniques. Field Manual 24-18. Washington, DC: Department of the Army.

USASC&FG. (April 1975) How to get out of a jam. TC 32-11. Washington, DC: Department of the Army.

USASC&FG. (September 1976) Combat communications (how to fight). Field Manual 24-1 (HTF). Washington, DC: Department of the Army.

USASC&FG. (March 1977) Combat communications within the division (how to fight). Field Manual 11-50 (HTF). Washington, DC: Department of the Army.

USASC&FG. (June 1977) Electronic counter-countermeasures (ECCM): Procedures for the communicator. TC 32-05-2PT. Washington, DC: Department of the Army.

USASC&FG. (September 1977) Communications in a 'come as you are' war. TC 24-18. Washington, DC: Department of the Army.

USASC&FG. (April 1978) Communications-electronics: Operations, orders, records and reports. Field Manual 24-16. Washington, DC: Department of the Army.

USASC&FG. (May 1982) (O) Communications-electronics operation instructions, the CEOI. TC 24-1. Washington, DC: Department of the Army.

USATSC. (September 1982) Tactical engagement simulations with MILES. TC 25-6. Washington, DC: Department of the Army.

Voiers, W.D. (October 1977), *Antecedents of speech acceptability.* Paper presented at the International Symposium on Measurements in Telecommunications, Lannion, France.

Ward, J.S., & Fooks, N.I. (December 1965) *Development of improved rifle squad tactical and patrolling programs for the light weapons infantryman* (HumRRO Technical Report 65-16). Alexandria, VA: Human Resources Research Organization.

Williamson, J.R.(June 1980) *The effects of Soviet Army communications jamming on the AIM Division Signal Battalion* (Thesis for Master of Military Art and Science). Fort Leavenworth, KS: U.S. Army Command and General Staff College.

Woodson, W.E., & Connover, D.W. (1964), *Human engineering guide for equipment designers.* Berkeley, California: University of California Press.

Wilson, S.H., Goodwin, N.C., Bersoff, E.H., & Thomas, N.M. (March 1982) *Military message experiment Final Report—Vol. 1, Executive Summary* (NRL Memorandum Report 4454). Washington, DC: Naval Research Laboratory.

Zipf, G.K. (1935) *The psycho-biology of language.* Boston, MA: Houghton Mifflin.

END

FILMED

8 84

DTIC